Multidim Syst Sign Process (2007) 18:191–248
DOI 10.1007/s11045-006-0011-y

ORIGINAL ARTICLE

Multivariable backward-shift-invariant subspaces and observability operators

Joseph A. Ball · Vladimir Bolotnikov · Quanlei Fang

Received: 29 January 2006 / Revised: 3 September 2006 /
Accepted: 7 September 2006 / Published online: 14 February 2007

Abstract It is well known that subspaces of the Hardy space over the unit disk which are invariant under the backward shift occur as the image of an observability operator associated with a discrete-time linear system with stable state-dynamics, as well as the functional-model space for a Hilbert space contraction operator. We discuss two multivariable extensions of this structure, where the classical Hardy space is replaced by (1) the Fock space of formal power series in a collection of d noncommuting indeterminates with norm-square-summable vector coefficients, and (2) the reproducing kernel Hilbert space (often now called the Arveson space) over the unit ball in $\mathbb{C}^d$ with reproducing kernel $k(\lambda, \zeta) = 1/(1 - \langle \lambda, \zeta \rangle)$ ($\lambda, \zeta \in \mathbb{C}^d$ with $\|\lambda\|, \|\zeta\| < 1$). In the first case, the associated linear system is of noncommutative Fornasini–Marchesini type with evolution along a free semigroup with d generators, while in the second case the linear system is a standard (commutative) Fornasini–Marchesini-type system with evolution along the integer lattice $\mathbb{Z}^d$. An abelianization map (or symmetrization of the Fock space) links the first case with the second. The second case has special features depending on whether the operator-tuple defining the state dynamics is commutative or not. The paper focuses on multidimensional state-output linear systems and the associated observability operators; followup papers Ball, Bollotnikov, and Fang (2007a, 2007b) use the results here to extend the analysis to represent observability-operator ranges as reproducing kernel Hilbert spaces with reproducing kernels constructed from the transfer function of a conservative multidimensional (noncommutative or commutative) input-state-output linear system.

J. A. Ball · Q. Fang
Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123, USA
e-mail: ball@math.vt.edu

V. Bolotnikov (✉)
Department of Mathematics, The College of William and Mary, Williamsburg, VA 23187-8795, USA
e-mail: vladi@math.wm.edu

Q. Fang
e-mail: qlfang@math.vt.edu

Keywords Operator valued functions · Schur multiplier

Mathematics Subject Classification 1991 47A57

1 Introduction

For $\mathcal{U}$ and $\mathcal{Y}$ any pair of Hilbert spaces, we use the notation $\mathcal{L}(\mathcal{U},\mathcal{Y})$ to denote the space of bounded, linear operators from $\mathcal{U}$ to $\mathcal{Y}$. For $\mathcal{X}$ a single Hilbert space, we shorten the notation $\mathcal{L}(\mathcal{X},\mathcal{X})$ to $\mathcal{L}(\mathcal{X})$. Let $\mathcal{X}$, $\mathcal{U}$, and $\mathcal{Y}$ be Hilbert spaces, let $A \in \mathcal{L}(\mathcal{X})$, $B \in \mathcal{L}(\mathcal{U},\mathcal{X})$, $C \in \mathcal{L}(\mathcal{X},\mathcal{Y})$, and $D \in \mathcal{L}(\mathcal{U},\mathcal{Y})$ be bounded linear operators, and let us consider the associated discrete-time linear time-invariant system

$$\begin{aligned} x(n+1) &= Ax(n) + Bu(n), \\ y(n) &= Cx(n) + Du(n) \end{aligned} \tag{1.1}$$

with $x(n)$ taking values in the *state space* $\mathcal{X}$, $u(n)$ taking values in the *input-space* $\mathcal{U}$ and $y(n)$ taking values in the *output-space* $\mathcal{Y}$. If we let the system evolve on the non-negative integers $n \in \mathbb{Z}_+$, then the whole trajectory $\{u(n), x(n), y(n)\}_{n\in\mathbb{Z}_+}$ is determined from the input signal $\{u(n)\}_{n\in\mathbb{Z}_+}$ and the initial state $x(0)$ according to the formulas

$$\begin{aligned} x(n) &= A^n x(0) + \sum_{k=0}^{n-1} A^{n-1-k} Bu(k), \\ y(n) &= CA^n x(0) + \sum_{k=0}^{n-1} CA^{n-1-k} Bu(k) + Du(n) \\ &= [\mathcal{O}_{C,A} x(0)]_n + \sum_{k=0}^{n-1} CA^{n-1-k} Bu(k) + Du(n), \end{aligned} \tag{1.2}$$

where $\mathcal{O}_{C,A}$ denotes the so-called *observability operator*

$$\mathcal{O}_{C,A}\colon x \mapsto \left\{CA^n x\right\}_{n\in\mathbb{Z}_+}.$$

If we introduce the Z-transform

$$\{f(n)\}_{n\in\mathbb{Z}_+} \mapsto \widehat{f}(\lambda) = \sum_{n=0}^{\infty} f(n)\lambda^n \tag{1.3}$$

the Z-transformed version of the system-trajectory formulas (1.2) become

$$\begin{aligned} \widehat{x}(\lambda) &= (I-\lambda A)^{-1} x(0) + \lambda (I-\lambda A)^{-1} B\widehat{u}(\lambda), \\ \widehat{y}(\lambda) &= C(I-\lambda A)^{-1} x(0) + [D + \lambda C(I-\lambda A)^{-1} B]\widehat{u}(\lambda) \\ &= \widehat{\mathcal{O}}_{C,A} x(0) + T_\Sigma(z)\widehat{u}(\lambda), \end{aligned} \tag{1.4}$$

where

$$\widehat{\mathcal{O}}_{C,A}\colon\; x \mapsto \sum_{n=0}^{\infty} (CA^n x)\,\lambda^n = C(I-\lambda A)^{-1} x \tag{1.5}$$

is the Z-transformed version of the observability operator and where

$$T_\Sigma(\lambda) = D + \lambda C(I - \lambda A)^{-1}B$$

is the *transfer function* of the system Σ given by (1.1). In particular, if the input signal $\{u(n)\}_{n\in\mathbb{Z}_+}$ is taken to be zero, the resulting output $\{y(n)\}_{n\in\mathbb{Z}_+}$ is given by $y = \mathcal{O}_{C,A}x(0)$. In case $\mathcal{O}_{C,A}$ is bounded as an operator from $\mathcal{X}$ into $\ell^2_{\mathcal{Y}} := \ell^2 \otimes \mathcal{Y}$ (here ℓ^2 is the space of square-summable complex sequences indexed by the non-negative integers $\mathbb{Z}_+$), we say that the pair (C, A) is *output-stable*. It is convenient to represent $\mathcal{O}_{C,A}$ in the output-stable case in the matrix form

$$\mathcal{O}_{C,A} = \mathrm{col}_{n\in\mathbb{Z}_+}\left[CA^n\right]\colon\ \mathcal{X} \to \ell^2_{\mathcal{Y}}.$$

Since the *Z-transform* (1.3) maps $\ell^2_{\mathcal{Y}}$ unitarily onto $H^2_{\mathcal{Y}} := H^2 \otimes \mathcal{Y}$, where H^2, the image of ℓ^2 under the Z-transform, is the space of analytic functions on the unit disk with modulus-square-summable sequence of Taylor coefficients:

$$H^2 = \left\{ f(\lambda) = \sum_{n=0}^{\infty} f_n\lambda^n\colon\ \sum_{n=0}^{\infty} |f_n|^2 < \infty \right\}$$

the output stability of (C, A) is equivalent to the Z-transformed version of the observability operator (1.5) being bounded as an operator from $\mathcal{X}$ into $H^2_{\mathcal{Y}}$. It is readily seen that $\widehat{\mathcal{O}}_{C,A}x = \widehat{\mathcal{O}_{C,A}x}$.

If (C, A) is output-stable, then the *observability gramian*

$$\mathcal{G}_{C,A} := (\mathcal{O}_{C,A})^*\mathcal{O}_{C,A} = (\widehat{\mathcal{O}}_{C,A})^*\widehat{\mathcal{O}}_{C,A}$$

is bounded on $\mathcal{X}$ and can be represented via the series

$$\mathcal{G}_{C,A} = \sum_{n=0}^{\infty} A^{*n}C^*CA^n \tag{1.6}$$

converging in the strong operator topology. The following result gives a summary of well-known connections between output stability, observability gramians and solutions of associated Stein equations and inequalities.

Theorem 1.1 *Let (C, A) be a pair of operators with $C\colon \mathcal{X} \to \mathcal{Y}$ and $A\colon \mathcal{X} \to \mathcal{X}$. Then:*

(1) *The pair (C, A) is output-stable if and only if the Stein inequality*

$$H - A^*HA \geq C^*C \tag{1.7}$$

has a positive semidefinite solution $H \in \mathcal{L}(\mathcal{X})$.

(2) *If (C, A) is output-stable, then the observability gramian $\mathcal{G}_{C,A}$ satisfies the Stein equality*

$$H - A^*HA = C^*C \tag{1.8}$$

and is the minimal positive semidefinite solution of the Stein inequality (1.7).

(3) *There is a unique positive semidefinite solution of the Stein equality* (1.8) *if A is* strongly stable, *i.e., powers A^n of A tend to zero in the strong operator topology of $\mathcal{L}(\mathcal{X})$. If A is a contraction operator, then the positive semidefinite solution of the Stein equation* (1.8) *is unique if and only if A is strongly stable.*

A pair (C, A) is called *observable* if the operator $\mathcal{O}_{C,A}$ (equivalently, $\widehat{\mathcal{O}}_{C,A}, \mathcal{G}_{C,A}$) is injective. This property means that a state space vector $x \in \mathcal{X}$ is uniquely recovered from the output string $\{y(n)\}_{n=0}^{\infty}$ generated by running the system (1.1) with the zero input string and the initial condition $x(0) = x$. A pair (C, A) is called *exactly observable* if $\mathcal{O}_{C,A}$ (equivalently, $\mathcal{G}_{C,A}$) is bounded and bounded from below.

Associated with an output-stable pair (C, A) is the range of the observability operator

$$\operatorname{Ran} \widehat{\mathcal{O}}_{C,A} = \{C(I - zA)^{-1}x \colon x \in \mathcal{X}\}.$$

The following theorem summarizes the connections between such ranges and backward-shift-invariant subspaces of $H^2_{\mathcal{Y}}$.

Theorem 1.2 *Suppose that (C, A) is an output-stable pair. Then:*

(1) *The linear manifold* $\operatorname{Ran} \widehat{\mathcal{O}}_{C,A}$ *is invariant under the backward shift operator*

$$S^{*}\colon f(\lambda) \to \frac{f(\lambda) - f(0)}{\lambda}. \tag{1.9}$$

(2) *Let $H \geq 0$ be a solution of the Stein inequality* (1.7) *and let $\mathcal{X}'$ be the completion of $\mathcal{X}$ with inner product $\|[x]\|^2_{\mathcal{X}'} = \langle Hx, x\rangle_{\mathcal{X}}$ (where $[x]$ denotes the equivalence class modulo* $\operatorname{Ker} H$ *generated by x). Then A and C extend to define bounded operators $A'\colon \mathcal{X}' \to \mathcal{X}'$ and $C'\colon \mathcal{X}' \to \mathcal{Y}$ and the observability operator $\widehat{\mathcal{O}}_{C,A}$ extends to define a contraction operator $\widehat{\mathcal{O}}_{C',A'}$ from $\mathcal{X}'$ into $H^2_{\mathcal{Y}}$. Moreover, $\widehat{\mathcal{O}}_{C',A'}\colon \mathcal{X}' \to H^2_{\mathcal{Y}}$ is an isometry if and only if H satisfies the Stein equation* (1.8) *and A' is strongly stable, i.e.,*

$$\langle HA^n x, A^n x\rangle \to 0 \quad \textit{for all} \;\; x \in \mathcal{X}.$$

(3) *If the linear manifold $\mathcal{M} :=$* $\operatorname{Ran} \widehat{\mathcal{O}}_{C,A}$ *is given the lifted norm*

$$\|\widehat{\mathcal{O}}_{C,A}x\|^2_{\mathcal{M}} = \inf_{y \in \mathcal{X}\colon \mathcal{O}_{C,A}y = \mathcal{O}_{C,A}x} \langle Hy, y\rangle_{\mathcal{X}},$$

then

(a) *$\mathcal{M}$ can be completed to $\mathcal{M}' =$* $\operatorname{Ran} \mathcal{O}_{C',A'}$ *with contractive inclusion in $H^2_{\mathcal{Y}}$:*

$$\|f\|^2_{H^2_{\mathcal{Y}}} \leq \|f\|^2_{\mathcal{M}'} \quad \textit{for all}\, f \in \mathcal{M}'.$$

Furthermore, $\mathcal{M}'$ is isometrically equal to the reproducing kernel Hilbert space with reproducing kernel $K_{C,A;H}$ given by

$$K_{C,A;H}(\lambda, \zeta) = C(I - \lambda A)^{-1} H (I - \overline{\zeta} A^{*})^{-1} C^{*}. \tag{1.10}$$

(b) *The following difference-quotient inequality is satisfied*

$$\|S^{*}f\|^2_{\mathcal{M}} \leq \|f\|^2_{\mathcal{M}} - \|f(0)\|^2_{\mathcal{Y}} \quad \textit{for all} \;\; f \in \mathcal{M} \tag{1.11}$$

and moreover, if the Stein equality (1.8) *holds, then* (1.11) *holds with equality.*

(4) *Conversely, if $\mathcal{M}$ is a Hilbert space contractively included in $H^2_{\mathcal{Y}}$ which is invariant under S^{*} and for which the difference-quotient inequality* (1.11) *holds, then there is a contractive pair (C, A) (i.e.,* (1.7) *holds with $H = I_{\mathcal{X}}$) such that $\mathcal{M} = \mathcal{H}(K_{C,A;I}) =$* $\operatorname{Ran} \mathcal{O}_{C,A}$ *isometrically. In case* (1.11) *holds with equality, then (C, A) can be taken to be isometric.*

Results of the type in Theorem 1.1 are the basis for the Lyapunov-function approach to stability analysis in system theory; there are far-reaching generalizations to non-linear and time-varying settings which are far afield from our main interests here. The goal of characterizing subspaces of H^2 of the form $\mathcal{H}(K_{C,A;H})$ (especially in a finite-dimensional context) was a key feature in the approach to Nevanlinna–Pick interpolation developed by Dym (1989).

In this paper, we present the analogues of Theorems 1.1 and 1.2 for the two related multivariable settings: (1) the case where the Hardy space H^2 on the unit disk is replaced by the Fock space $H^2_{\mathcal{Y}}(\mathcal{F}_d)$, and (2) the case where H^2 is replaced by the vector-valued Arveson reproducing kernel Hilbert space $\mathcal{H}_{\mathcal{Y}}(k_d)$.

To define the Fock space, we let $\mathcal{F}_d$ denote the free semigroup on the set $\{1,\dots,d\}$ of the first d natural numbers and then let $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ consist of the space of all formal power series $\sum_{v\in\mathcal{F}_d} f_v z^v$ in d noncommuting indeterminates $z=(z_1,\dots,z_d)$ with coefficients f_v in a coefficient Hilbert space $\mathcal{Y}$ which are square-summable in norm: $\sum_{v\in\mathcal{F}_d}\|f_v\|^2_{\mathcal{Y}}<\infty$. Here we write $z^v=z_{i_N}z_{i_{N-1}}\dots z_{i_1}$ if $v=i_N i_{N-1}\dots i_1\in\mathcal{F}_d$. The shift operator S: $f(\lambda)\mapsto\lambda f(\lambda)$ acting on the Hardy space H^2 is replaced by the noncommuting d-tuple $\mathbf{S}=(S_1,\dots,S_d)$ on $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ given by

$$S_j\colon f(z)\mapsto f(z)z_j \quad \text{for} \quad j=1,\dots,d. \tag{1.12}$$

The system (1.1) is replaced by a noncommutative multidimensional input-state-output system of the form

$$\begin{array}{rcl} x(1v) &=& A_1x(v)+B_1u(v),\\ \vdots & \vdots & \vdots\\ x(dv) &=& A_dx(v)+B_du(v),\\ y(v) &=& Cx(v)+Du(v). \end{array} \tag{1.13}$$

Here the system evolves along the free semigroup $\mathcal{F}_d$, and, for each $v\in\mathcal{F}_d$, the state vector $x(v)$, input signal $u(v)$ and output signal $y(v)$ take values in the *state space* $\mathcal{X}$, *input space* $\mathcal{U}$, and *output space* $\mathcal{Y}$, and the *system matrix* U has the form

$$U=\begin{bmatrix} A & B\\ C & D\end{bmatrix}=\begin{bmatrix} A_1 & B_1\\ \vdots & \vdots\\ A_d & B_d\\ C & D\end{bmatrix}\colon\begin{bmatrix}\mathcal{X}\\ \mathcal{U}\end{bmatrix}\to\begin{bmatrix}\mathcal{X}\\ \vdots\\ \mathcal{X}\\ \mathcal{Y}\end{bmatrix}. \tag{1.14}$$

Such systems were introduced in Ball and Vinnikov (2005) and with further elaboration in Ball, Groenewald, and Malakorn (2005, 2006), following Ball, Groenewald, and Malakorn (2005), we call this type of system a *noncommutative Fornasini–Marchesini linear system*. The observability operator associated with an output map $C\colon\mathcal{X}\to\mathcal{Y}$ and a d-tuple $\mathbf{A}=(A_1,\dots,A_d)$ of not necessarily commuting operators on a Hilbert space $\mathcal{X}$, expressed in "frequency-domain" coordinates, takes the form

$$\widehat{\mathcal{O}}_{C,\mathbf{A}}\colon x\mapsto C(I-z_1A_1-\cdots-z_dA_d)^{-1}x.$$

For the particular case where $\mathbf{A}$ is a row contraction and

$$C=(I-A_1^*A_1-\cdots-A_d^*A_d)^{1/2}$$

with $\mathcal{Y}$ taken to be equal to the closure of the range of C, this operator appears already in work of Popescu (1999) under the term "Poisson kernel" and as the adjoint of the key operator L used in many constructions in the paper of Arveson (2000). Reproducing kernel Hilbert spaces consisting of formal power series were developed in a systematic way in Ball and Vinnikov (2003). Such spaces already appear (although not quite in our notation) in the Sz.-Nagy–Foiaş model theory for row contractions developed by Popescu (1989a–c). We shall see that Theorems 1.1 and 1.2 extend in a natural way to this setting, where the observability gramian (1.6) in the statement of Theorem 1.1 is replaced with the multivariable observability gramian

$$\mathcal{G}_{C,\mathbf{A}} = \widehat{\mathcal{O}}_{C,\mathbf{A}}^{*} \widehat{\mathcal{O}}_{C,\mathbf{A}} = \sum_{v \in \mathcal{F}_d} A^{v*} C^* C A^{v} \tag{1.15}$$

(here we set $A^v = A_{i_N} \dots A_{i_1}$ if $v = i_N \dots i_1 \in \mathcal{F}_d$), where the backward shift S^* (1.9) in the statement of Theorem 1.2 is replaced by the d-tuple $\mathbf{S}^* = (S_1^*, \dots, S_d^*)$ of adjoints of the shift operators S_j in (1.12), and where the positive kernel (1.10) becomes the kernel

$$K_{C,\mathbf{A},H}(z,w) = C(I - z_1 A_1 - \cdots - z_d A_d)^{-1} H (I - w_1 A_1^* - \cdots - w_d A_d^*)^{-1} C^* \tag{1.16}$$

in two sets $z = (z_1, \dots, z_d)$ and $w = (w_1, \dots, w_d)$ of noncommuting indeterminates (see Theorems 2.2 and 2.8 below).

In the second Arveson-space setting, the Hardy space H^2 over the unit disk is replaced by the so-called Arveson space, the reproducing kernel Hilbert space $\mathcal{H}(k_d)$ over the unit ball $\mathbb{B}^d$ in complex d-dimensional space $\mathbb{C}^d$ based on the reproducing kernel function

$$k_d(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = \frac{1}{1 - \langle \boldsymbol{\lambda}, \boldsymbol{\zeta} \rangle_{\mathbb{C}^d}} \quad \text{for } \boldsymbol{\lambda}, \boldsymbol{\zeta} \in \mathbb{B}^d$$

and the classical Hardy-space shift $f(\lambda) \mapsto \lambda \cdot f(\lambda)$ is replaced by the d-tuple of Arveson shift operators $\mathbf{M}_{\boldsymbol{\lambda}} = (M_{\lambda_1}, \dots, M_{\lambda_d})$ where

$$M_{\lambda_j} \colon f(\boldsymbol{\lambda}) \mapsto \lambda_j f(\boldsymbol{\lambda}) \quad \text{for } f \in \mathcal{H}(k_d) \tag{1.17}$$

(Arveson, 2000; Drury, 1978). In this case the underlying system evolves along the integer lattice $\mathbb{Z}_+^d = \{\mathbf{n} = (n_1, \dots, n_d) \colon n_j \in \mathbb{Z}_+\}$ and has the form of what we call a *(commutative) Fornasini–Marchesini system*

$$\begin{aligned} x(\mathbf{n}) &= A_1 x(\mathbf{n} - e_1) + \cdots + A_d x(\mathbf{n} - e_d) \\ &\quad + B_1 u(\mathbf{n} - e_1) + \cdots + B_d u(\mathbf{n} - e_d), \\ y(\mathbf{n}) &= C x(\mathbf{n}) + D u(\mathbf{n}). \end{aligned} \tag{1.18}$$

Here and in what follows, e_j denotes the element in $\mathbb{Z}_+^d$ having the jth partial index equal to one and all other partial indices equal to zero:

$$e_j = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{Z}_+^d. \tag{1.19}$$

Thus, the system matrix U has the same form (1.14) as for the noncommutative setting but the domain for all the signals and the system evolution is the integer lattice $\mathbb{Z}_+^d$ rather than the free semigroup $\mathcal{F}_d$ and the associated "frequency-domain" objects are functions or formal power series in the commuting variables $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_d)$ rather than in the noncommuting indeterminates $z = (z_1, \dots, z_d)$. In Sect. 3, we show how the Arveson space $\mathcal{H}(k_d) \otimes \mathcal{Y}$ and this Fornasini–Marchesini linear system can be derived as an abelianization (sometimes also called *symmetrization*) of the noncommutative Fock space $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ and of the noncommutative Fornasini–Marchesini linear system, respectively; while it is well known that the Arveson space is a symmetrization of the Fock space and that the multiplier algebra on the Arveson space is the image under a completely positive map acting on the noncommutative multiplier algebra on the Fock space (Arias & Popescu, 2000; Arveson, 1998; Davidson & Pitts, 1998; Davidson, 2001, and Popescu (2006; to appear in Proc. Amer. Math. Soc.) for a recent, more general systematic framework), our extension of these ideas to the underlying system theory appears to be new. The observability operator, as in the noncommutative setting, is associated with a so-called output pair $(C, \mathbf{A})$ but now has the form

$$\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} \colon x \mapsto C(I - \lambda_1 A_1 - \cdots - \lambda_d A_d)^{-1} x,$$

where the variables $\lambda_1, \ldots, \lambda_d$ commute and the *abelianized* observability gramian $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = (\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}})^* \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ has an infinite-series representation more complicated than the second expression in (1.15) (see Eq. 3.10 below). In case the operator d-tuple $\mathbf{A} = (A_1, \ldots, A_d)$ is *commutative* (so $A_i A_j = A_j A_i$ for all $1 \le i, j \le d$), $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}}$ (see Proposition 3.3 below), and Theorem 1.2 has a natural analogue for this setting, with the abelianized multivariable observability gramian $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}}$ playing the role of the observability gramian in Theorem 1.1, with the operator d-tuple $\mathbf{M}^*_{\boldsymbol{\lambda}} = (M^*_{\lambda_1}, \ldots, M^*_{\lambda_d})$ in place of the backward shift S^* (1.9) in Theorem 1.2, and with kernel (1.10) now taken to be the multivariable positive kernel

$$K^{\mathbf{a}}_{C,\mathbf{A};H}(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = C(I - \lambda_1 A_1 - \cdots - \lambda_d A_d)^{-1} H (I - \overline{\zeta_1} A_1^* - \cdots - \overline{\zeta_d} A_d^*)^{-1} C^* \quad (1.20)$$

(see Theorems 3.14–3.16 below). In the general case where the operator d-tuple $\mathbf{A} = (A_1, \ldots, A_d)$ is not assumed to be commutative, there is no characterization of the abelianized observability gramian as a minimal solution of a generalized Stein equation analogous to the classical case given in Theorem 1.1, but there still is a somewhat more implicit analogue of Theorem 1.2, where the backward shift S^* (1.9) in Theorem 1.2 is replaced by a solution of the so-called *Gleason problem* (see Theorems 3.20 and 3.21 below). The Gleason problem originates in the work of Gleason (1964) and Henkin (1971) and has been studied in the context of the Arveson space (with various formulas for the solution) in Alpay and Dubi (2005) with an application to realization questions in Alpay, Dijksma, and Rovnyak (2003). Our analogue of Theorem 1.2 for the Arveson space for the case of commutative d-tuple $\mathbf{A}$ has already been given in Bolotnikov and Rodman (2002) (with a more general power-series setting worked out in Bolotnikov & Rodman, 2004) for the finite-dimensional case.

We also give various numerical examples (constructed with the aid of the software program MATHEMATICA) to illustrate how $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ and $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ can have divergent properties when $\mathbf{A}$ is not commutative (see Examples 3.4, 3.9, and 3.11 below).

Backward-shift-invariant subspaces for the classical setting have been used for some time in the operator-theory literature as the model space for a more general (abstractly defined) Hilbert-space contraction operator (de Branges & Rovnyak, 1966; Sz.-Nagy & Foiaş, 1970); connections of this work with linear system theory were only realized later (see, e.g., Helton, 1972/1973, 1974). Our results develop the structure of such model spaces for the case of operator-tuples and therefore are of interest from the point of view of multivariable operator theory. We find it satisfying that these model spaces in turn tie in with the theory of multidimensional linear systems in much the same way (but with some surprises) as in the classical case.

As applications of the ideas, we obtain new system-theoretic derivations of the Beurling–Lax representation theorem for shift invariant subspaces in both the noncommutative and commutative settings; the result for the noncommutative setting is due originally to Popescu (1989d) and for the commutative setting to Arveson (1998) and McCullough and Trent (2000). We also indicate connections with dilation theory and the von Neumann inequality for these settings (see Arveson, (1998); Drury, 1978; Popescu, 1991,1999).

Closely related to the kernels $K_{C,\mathbf{A}}(z,w)$ and $K^{\mathbf{a}}_{C,\mathbf{A}}(\boldsymbol{\lambda},\boldsymbol{\zeta})$ (given by (1.16) and (1.20) with H normalized to be the identity operator) are kernels of de Branges–Rovnyak type (see de Branges & Rovnyak, 1966 for the classical case)

$$K_S(z,w) = k_{\mathrm{Sz}}(z,w) - S(z)k_{\mathrm{Sz}}(z,w)S(w)^*, \qquad K^{\mathbf{a}}_S(\boldsymbol{\lambda},\boldsymbol{\zeta}) = \frac{I - S(\boldsymbol{\lambda})S(\boldsymbol{\zeta})^*}{1 - \langle \boldsymbol{\lambda},\boldsymbol{\zeta}\rangle}$$

(where $z = (z_1,\ldots,z_d)$ and $w = (w_1,\ldots,w_d)$ are two sets of noncommuting indeterminates with $k_{\mathrm{Sz}}(z,w) = \sum_{v\in\mathcal{F}_d} z^v w^{v^\top}$ equal to the *noncommutative Szegö kernel* while $\boldsymbol{\lambda} = (\lambda_1,\ldots,\lambda_d)$ and $\boldsymbol{\zeta} = (\zeta_1,\ldots,\zeta_d)$ are two sets of commuting variables) for respective reproducing kernel Hilbert spaces $\mathcal{H}(K_S)$, $\mathcal{H}(K^{\mathbf{a}}_S)$ in the respective noncommutative and commutative settings. In this situation (where K_S and $K^{\mathbf{a}}_S$ are positive kernels in noncommuting and commuting variables, respectively), the respective power series

$$S(z) = \sum_{v\in\mathcal{F}_d} S_v z^v, \qquad S(\lambda) = \sum_{\mathbf{n}\in\mathbb{Z}^d_+} S_{\mathbf{n}}\boldsymbol{\lambda}^{\mathbf{n}}$$

are *contractive multipliers*, i.e., the respective multiplication operators

$$M_S\colon f(z)\mapsto S(z)\cdot f(z), \qquad M_S\colon f(\boldsymbol{\lambda})\mapsto S(\boldsymbol{\lambda})\cdot f(\boldsymbol{\lambda})$$

are bounded from $H^2_{\mathcal{U}}(\mathcal{F}_d)$ into $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ and from $\mathcal{H}_{\mathcal{U}}(k_d)$ into $\mathcal{H}_{\mathcal{Y}}(k_d)$, respectively, with norm at most 1. A particular issue is the construction of operators

$$B_j\colon \mathcal{U}\to\mathcal{X}, \quad \text{for } j = 1,\ldots,d \quad \text{and} \quad D\colon \mathcal{U}\to\mathcal{Y}$$

for some input space $\mathcal{U}$ so that

$$\begin{aligned} S(z) &= D + C(I - z_1A_1 - \cdots - z_dA_d)^{-1}(z_1B_1 + \cdots + z_dB_d),\\ S(\boldsymbol{\lambda}) &= D + C(I - \lambda_1A_1 - \cdots - \lambda_dA_d)^{-1}(\lambda_1B_1 + \cdots + \lambda_dB_d)\end{aligned}$$

satisfy

$$K_{C,\mathbf{A}}(z,w) = K_S(z,w), \qquad K^{\mathbf{a}}_{C,\mathbf{A}}(\boldsymbol{\lambda},\boldsymbol{\zeta}) = K^{\mathbf{a}}_S(\boldsymbol{\lambda},\boldsymbol{\zeta}).$$

With the resolution of this issue, then the results here lead directly to representations of backward-shift-invariant subspaces as reproducing kernel Hilbert spaces of the form $\mathcal{H}(K_S)$ and $\mathcal{H}(K^{\mathbf{a}}_S)$ for a Schur multiplier S in both the noncommutative and commutative settings as well as linear-fractional realizations for Beurling–Lax representers of shift-invariant subspaces for both the noncommutative (see Popescu, (1989d)) and commutative (see McCullough & Trent, (2000)) settings. We work out these issues for the commutative setting and for the noncommutative setting in Ball, Bolotnikov, and Fang (2007a) and Ball, Bolotnikov, and Fang (2007b) respectively.

The paper is organized as follows. After the present Introduction, Sect. 2 focuses on the noncommutative Fock space setting while Sect. 3 focuses on the Arveson-space setting. Section 2 is divided into Sect. 2.1 dealing with the connections between solutions of generalized Stein equations and strong stability of the state dynamics for noncommutative Fornasini–Marchesini systems and Sect. 2.2 dealing with characterizing ranges of observability operators as backward-shift-invariant subspaces of the Fock space with a certain reproducing-kernel-Hilbert-space structure. The first subsection (Sect. 3.1) of Sect. 3 deals with the less tractable issues parallel to the material in Sect. 2.1 of generalized Stein equations and stability for commutative Fornasini–Marchesini systems and also presents the abelianization map giving the connection between noncommutative and commutative Fornasini–Marchesini systems. The second subsection (Sect. 3.2) of Sect. 3, parallel to Sect. 2.2, discusses characterizations of observability-operator ranges for the case of a commutative Fornasini–Marchesini state-output system. The results are the most satisfying in case the operator-tuple $\mathbf{A}$ giving the state dynamics is commutative—these are collected in Sect. 3.2.1. The more implicit results for the case of noncommutative $\mathbf{A}$ are given in Sect. 3.2.2.

2 The Fock-space setting

2.1 Output stability and Stein equations: the noncommutative case

For d a positive integer, let $\mathcal{F}_d$ be the free semigroup $\mathcal{F}_d$ generated by the set of d letters $\{1, \ldots, d\}$. Elements of $\mathcal{F}_d$ are words of the form $i_N \ldots i_1$ where $i_\ell \in \{1, \ldots, d\}$ for each $\ell = 1, \ldots, N$ with multiplication given by concatenation. We also use $\emptyset$ to denote the empty word; this serves as the unit element for $\mathcal{F}_d$. For $v = i_N i_{N-1} \ldots i_1 \in \mathcal{F}_d$, we let $|v|$ denote the number N of letters in v and we let $v^\top := i_1 \ldots i_{N-1} i_N$ denote the *transpose* of v. We let $z = (z_1, \ldots, z_d)$ to be a collection of d formal noncommuting variables and let $\mathcal{Y}\langle\langle z\rangle\rangle$ denote the set of formal noncommutative series $\sum_{v \in \mathcal{F}_d} f_v z^v$ where $f_v \in \mathcal{Y}$ and where

$$z^v = z_{i_N} z_{i_{N-1}} \ldots z_{i_1}, \quad \text{if} \quad v = i_N i_{N-1} \ldots i_1. \tag{2.1}$$

The Fock space $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$ is defined as

$$\ell^2_{\mathcal{Y}}(\mathcal{F}_d) := \left\{ \{f_v\}_{v \in \mathcal{F}_d} \colon \sum_{v \in \mathcal{F}_d} \|f_v\|^2_{\mathcal{Y}} < \infty \right\}. \tag{2.2}$$

If we let χ_v be the characteristic function of the word v, so

$$\chi_v = \{\chi_v(v')\}_{v' \in \mathcal{F}_d}, \quad \text{where} \quad \chi_v(v') = \begin{cases} 1, & \text{if } v' = v, \\ 0, & \text{otherwise} \end{cases}$$

and we let $\mathcal{B}_{\mathcal{Y}}$ be an orthonormal basis for $\mathcal{Y}$, then $\{\chi_v y_i \colon v \in \mathcal{F}_d, y_i \in \mathcal{B}_{\mathcal{Y}}\}$ is an orthonormal basis for $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$. The space $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$ can be identified as the tensor product $\ell^2(\mathcal{F}_d) \otimes \mathcal{Y}$ and is mapped unitarily onto the space

$$H^2_{\mathcal{Y}}(\mathcal{F}_d) = \left\{ \sum_{v \in \mathcal{F}_d} f_v z^v \in \mathcal{Y}\langle\langle z\rangle\rangle \colon \sum_{v \in \mathcal{F}_d} \|f_v\|^2_{\mathcal{Y}} < \infty \right\} \tag{2.3}$$

by the noncommutative Z-transform

$$\{f_v\}_{v \in \mathcal{F}_d} \mapsto f^\wedge(z) = \sum_{v \in \mathcal{F}_d} f_v z^v \tag{2.4}$$

with the monomials z^v playing the role of the basis vectors χ_v.

The noncommutative multidimensional analogue of the system (1.1) is the system with evolution along the free semigroup $\mathcal{F}_d$ given by (1.13). Upon running the system (1.13) with the zero input string $u(v) = 0$ for $v \in \mathcal{F}_d$ and a fixed initial condition $x(\emptyset) = x \in \mathcal{X}$ we get

$$y(v) = C\mathbf{A}^v x + \sum_{v'',v' \in \mathcal{F}_d, j \in \{1,\ldots,d\}:\ v''jv'=v} C\mathbf{A}^{v''} B_j u(v'). \tag{2.5}$$

Here we extend the noncommutative functional calculus (2.1) from noncommuting indeterminates $z = (z_1, \ldots, z_d)$ to a d-tuple of operators $\mathbf{A} = (A_1, \ldots, A_d)$; we use the notation

$$\mathbf{A}^v = A_{i_N} A_{i_{N-1}} \ldots A_{i_1}, \quad \text{if} \quad v = i_N i_{N-1} \ldots i_1 \in \mathcal{F}_d,$$

where the multiplication is now operator composition. Application of the formal noncommutative Z-transform (2.4) then gives

$$\widehat{y}(z) = C(I - Z(z)A)^{-1} x(\emptyset) + T_\Sigma(z)\widehat{u}(z), \tag{2.6}$$

where the formal power series $T_\Sigma(z)$ (by definition equal to the *transfer function* of the system (1.13)) is given by

$$T_\Sigma(z) = D + C(I - Z(z)A)^{-1} Z(z)B,$$

where we have set

$$Z(z) = \begin{bmatrix} z_1 & \cdots & z_d \end{bmatrix} \otimes I_{\mathcal{X}}, \quad A = \begin{bmatrix} A_1 \\ \vdots \\ A_d \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_d \end{bmatrix}. \tag{2.7}$$

For details see Ball and Vinnikov (2005) or, for a more general setting of structured noncommutative multidimensional systems, see Ball, Sadosky, and Vinnikov (2005).

In analogy to the classical case, the system (1.13) is called *output-stable* (and in this case we will say that the pair $(C, \mathbf{A})$ is output-stable) if the output string $\{y(v)\}_{v \in \mathcal{F}_d}$, defined as in (2.5) but with the input string $\{u(v)\}_{v \in \mathcal{F}_d}$ assumed to be equal to 0, belongs to $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$ for every $x \in \mathcal{X}$ and the observability operator

$$\mathcal{O}_{C,\mathbf{A}}\colon\ x \mapsto \left\{C\mathbf{A}^v x\right\}_{v \in \mathcal{F}_d} \tag{2.8}$$

is bounded as an operator from $\mathcal{X}$ into $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$. The Z-transformed version of $\mathcal{O}_{C,\mathbf{A}}$ is

$$\widehat{\mathcal{O}}_{C,\mathbf{A}}\colon\ x \mapsto \sum_{v \in \mathcal{F}_d} (C\mathbf{A}^v x)\, z^v \in \mathcal{Y}\langle\langle z \rangle\rangle$$

and the following realization formula for $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ is immediate:

$$\widehat{\mathcal{O}}_{C,\mathbf{A}} x = C(I - Z(z)A)^{-1} x.$$

If $(C, \mathbf{A})$ is output-stable, then $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ maps $\mathcal{X}$ into $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ and is bounded. In this case it makes sense to introduce the *observability gramian*

$$\mathcal{G}_{C,\mathbf{A}} := (\mathcal{O}_{C,\mathbf{A}})^* \mathcal{O}_{C,\mathbf{A}} = (\widehat{\mathcal{O}}_{C,\mathbf{A}})^* \widehat{\mathcal{O}}_{C,\mathbf{A}} \tag{2.9}$$

and its representation in terms of strongly converging series

$$\mathcal{G}_{C,\mathbf{A}} = \sum_{v\in\mathcal{F}_d} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v \tag{2.10}$$

follows immediately by definition (2.8) of $\mathcal{O}_{C,\mathbf{A}}$ and the formula (2.2) for the norm in $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$. The second equality in (2.9) follows by definition of $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ and the formula (2.3) for the norm in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$.

Definition 2.1 A pair $(C, \mathbf{A})$ is called *observable* if $\mathcal{G}_{C,\mathbf{A}}$ is positive definite and *exactly observable* if $\mathcal{G}_{C,\mathbf{A}}$ is strictly positive definite. We say that the d-tuple of operators $\mathbf{A} = (A_1, \ldots, A_d)$ is *strongly stable* if

$$\lim_{N\to\infty} \sum_{v\in\mathcal{F}_d \colon |v|=N} \|\mathbf{A}^v x\|^2 \to 0 \quad \text{for all } \ x \in \mathcal{X}. \tag{2.11}$$

We mention that the term *pure* rather than *strongly stable* has been used in this context (see Arveson, 2000), but we prefer the present terminology so as to avoid confusion with the use of the term *pure* in the context of contractive operator-valued functions (see Sz.-Nagy & Foiaş, 1970).

In analogy with the classical case one can introduce the unobservable subspace

$$\operatorname{Ker} \mathcal{G}_{C,\mathbf{A}} = \operatorname{Ker} \mathcal{O}_{C,\mathbf{A}} = \bigcap_{v\in\mathcal{F}_d} \operatorname{Ker} C\mathbf{A}^v. \tag{2.12}$$

Thus, observability of $(C, \mathbf{A})$ means that $\operatorname{Ker} \mathcal{G}_{C,\mathbf{A}}$ is the zero subspace or that

$$C\mathbf{A}^v x = 0 \ \ (\forall v \in \mathcal{F}_d) \quad \Longrightarrow \quad x = 0. \tag{2.13}$$

The following is the noncommutative Fock-space counterpart to Theorem 1.1.

Theorem 2.2 *Let $\mathbf{A} = (A_1, \ldots, A_d) \in \mathcal{L}(\mathcal{X})^d$ and let $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then the pair $(C, \mathbf{A})$ is output-stable if and only if the (generalized) Stein inequality*

$$H - A_1^* H A_1 - \cdots - A_d^* H A_d \geq C^* C \tag{2.14}$$

has a positive semidefinite solution $H \in \mathcal{L}(\mathcal{X})$. In this case,

(1) *The observability gramian $\mathcal{G}_{C,\mathbf{A}}$ satisfies the generalized Stein equation*

$$H - A_1^* H A_1 - \cdots - A_d^* H A_d = C^* C \tag{2.15}$$

and is the minimal positive semidefinite solution of the generalized Stein inequality (2.14).

(2) *The positive semidefinite solution of the Stein equation* (2.15) *is unique if $\mathbf{A}$ is* strongly stable, *i.e.,* (2.11) *holds. Moreover, in case $\mathbf{A}$ is contractive in the sense that*

$$A_1^* A_1 + \cdots + A_d^* A_d \leq I_{\mathcal{X}}, \tag{2.16}$$

then the solution of the Stein equation (2.15) *is unique if and only if $\mathbf{A}$ is strongly stable.*

Proof Suppose first that $(C, \mathbf{A})$ is output-stable. Then for each $x \in \mathcal{X}$,

$$\{C\mathbf{A}^v x\}_{v\in\mathcal{F}_d} \in \ell^2_{\mathcal{Y}}(\mathcal{F}_d), \quad \text{i.e.,} \sum_{v\in\mathcal{F}_d} \|C\mathbf{A}^v x\|^2_{\mathcal{Y}} < \infty.$$

This has the consequence that the infinite series

$$\sum_{N=0}^{\infty} \sum_{v \in \mathcal{F}_d \colon |v|=N} \mathbf{A}^{v*} C^* C \mathbf{A}^v$$

converges in the strong operator topology to an operator $H \in \mathcal{L}(\mathcal{X})$ (in fact, $H = \mathcal{G}_{C,\mathbf{A}}$ is the observability gramian). From this infinite-series representation for $\mathcal{G}_{C,\mathbf{A}}$ it is easily verified that $\mathcal{G}_{C,\mathbf{A}}$ is positive semidefinite and satisfies the Stein equation (2.15) and hence also the Stein inequality (2.14).

Conversely, suppose that the Stein inequality (2.15) has a positive semidefinite solution H. We first claim that

$$H \geq \sum_{v \in \mathcal{F}_d \colon |v| \leq N} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v + \sum_{v \in \mathcal{F}_d \colon |v|=N+1} \mathbf{A}^{*v^\top} H \mathbf{A}^v \tag{2.17}$$

for each $N \in \mathbb{Z}_+$. For $N = 0$, (2.17) collapses to (2.14) which is given. Inductively assume that

$$H \geq \sum_{v \in \mathcal{F}_d \colon |v| < N} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v + \sum_{v \in \mathcal{F}_d \colon |v|=N} \mathbf{A}^{*v^\top} H \mathbf{A}^v.$$

Use the Stein inequality (2.14) to replace H on the right side by its lower bound $C^*C + A_1^* H A_1 + \cdots + A_d^* H A_d$ to get from this

$$H \geq \sum_{v \in \mathcal{F}_d \colon |v| < N} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v + \sum_{v \in \mathcal{F}_d \colon |v|=N+1} \mathbf{A}^{*v^\top} H \mathbf{A}^v + \sum_{v \in \mathcal{F}_d \colon |v|=N} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v,$$

which then simplifies to (2.17) as wanted.

We rewrite (2.17) in the form

$$\sum_{v \in \mathcal{F}_d \colon |v| < N} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v \leq H - \sum_{v \in \mathcal{F}_d \colon |v|=N} \mathbf{A}^{*v^\top} H \mathbf{A}^v \leq H. \tag{2.18}$$

By letting $N \to \infty$ in (2.18), we conclude that the left hand side sum converges (weakly and therefore, since all the terms are positive semidefinite, strongly) to a bounded positive semidefinite operator. By (2.10),

$$\lim_{N \to \infty} \sum_{v \in \mathcal{F}_d \colon |v| < N} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v = \sum_{v \in \mathcal{F}_d} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v = \mathcal{G}_{C,\mathbf{A}}$$

and passing to the limit in (2.18) as $N \to \infty$ gives $\mathcal{G}_{C,\mathbf{A}} \leq H$. In particular the operator $\mathcal{G}_{C,\mathbf{A}}$ is bounded (since H is) and therefore the pair $(C, \mathbf{A})$ is output-stable. □

Proof of (1) As observed in the proof of the first part of the theorem, from the infinite-series representation (2.10) it follows that $\mathcal{G}_{C,\mathbf{A}}$ satisfies the Stein equation (2.15). If H is any solution of the Stein inequality, the computation leading to (2.18) shows that H satisfies (2.18). By taking the limit as $N \to \infty$ we conclude that $\mathcal{G}_{C,\mathbf{A}} \leq H$ as asserted. □

Proof of (2) Suppose that $\mathbf{A}$ is strongly stable and that H solves the Stein equation (2.15). Then the proof of (2.17) shows that in this case (2.17) holds with equality:

$$H = \sum_{v \in \mathcal{F}_d \colon |v| \leq N} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v + \sum_{v \in \mathcal{F}_d \colon |v|=N+1} \mathbf{A}^{*v^\top} H \mathbf{A}^v \tag{2.19}$$

for each $N = 0, 1, 2, \ldots$ Taking the limit as $N \to \infty$ and using the stability assumption (2.11) we conclude that $H = \mathcal{G}_{C,\mathbf{A}}$.

For the converse direction we assume in addition that $\mathbf{A}$ is contractive (i.e., (2.16) holds). We prove the contrapositive: *if* $\mathbf{A}$ *does not satisfy the stability condition* (2.11)*, then the solution of the Stein equation* (2.15) *is not unique.* Assume therefore that $\mathbf{A}$ is not stable. By the assumption (2.16), the sequence of operators

$$\Delta_N = \sum_{v \in \mathcal{F}_d \colon |v| = N} \mathbf{A}^{*v^\top} \mathbf{A}^v, \ N = 1, 2, \ldots$$

is decreasing and bounded below and therefore has a strong limit Δ. Since $\mathbf{A}$ is assumed not to be stable, this limit Δ is not zero. However it is easily verified that

$$A_1^* \Delta_N A_1 + \cdots + A_d^* \Delta_N A_d = \Delta_{N+1}. \tag{2.20}$$

Taking limits in (2.20) gives that $\Delta = \lim_{N\to\infty} \Delta_N$ satisfies the homogeneous Stein equation

$$\Delta - A_1^* \Delta A_1 - \cdots - A_d^* \Delta A_d = 0.$$

We conclude that the solution of the Stein equation (2.15) cannot be unique. □

Particular cases of output pairs $(C, \mathbf{A})$ are the cases where $(C,\mathbf{A})$ is *contractive* (i.e., the Stein inequality (2.14) holds with $H = I_{\mathcal{X}}$) and where $(C,\mathbf{A})$ is *isometric* (i.e., the Stein equality (2.15) holds with $H = I_{\mathcal{X}}$). For these cases some additional observations can be made along the lines of Theorem 2.2.

Proposition 2.3

(1) *Suppose that* $(C, \mathbf{A})$ *is a contractive pair. Then* $(C,\mathbf{A})$ *is output-stable with* $\mathcal{G}_{C,\mathbf{A}} \leq I_{\mathcal{X}}$ *and the observability gramian* $\mathcal{G}_{C,\mathbf{A}}$ *is the unique positive semidefinite solution of the Stein equation* (2.15) *if and only if* $\mathbf{A}$ *is strongly stable.*
(2) *Suppose that* $(C, \mathbf{A})$ *is an isometric pair. Then* $(C,\mathbf{A})$ *is output-stable. Moreover* $H = I_{\mathcal{X}}$ *is the unique solution of the Stein equation* (2.15) *if and only if* $\mathbf{A}$ *is strongly stable. In this case* $\mathcal{O}_{C,\mathbf{A}}$ *is isometric and hence also* $(C,\mathbf{A})$ *is exactly observable.*

Proof Statement (1) immediately follows from statements in Theorem 2.2 combined with the observation that $(C,\mathbf{A})$ being a contractive pair implies that $\mathbf{A}$ is contractive.

The first two assertions in statement (2) follow in a similar way. As for the last assertion, for the case where $(C, \mathbf{A})$ is isometric, $I_{\mathcal{X}}$ is a solution of the Stein equation (2.15); for the situation where $\mathbf{A}$ is strongly stable, uniqueness implies that the observability gramian $\mathcal{G}_{C,\mathbf{A}} = I_{\mathcal{X}}$, i.e., that $\mathcal{O}_{C,\mathbf{A}}$ is isometric. Then also $(C, \mathbf{A})$ is exactly observable by definition. □

Remark 2.4 The converse of the last part of Proposition 2.3 does not hold even for the case $d = 1$. More precisely, *there exists an isometric pair of operators* (C, A) *such that* (C, A) *is observable but* A *is not strongly stable.*

An example necessarily requires that $\dim \mathcal{X} = \infty$. In the terminology of Sz.-Nagy and Foiaş (1970), it suffices to produce a *completely nonisometric* (*c.n.i.*) contraction operator A on a nontrivial Hilbert space $\mathcal{X}$ (so $\dim \mathcal{X} > 0$ and there is no nonzero-invariant subspace $\mathcal{M}$ for A such that $A|_{\mathcal{M}}$ is an isometry) in the class $C_{1\cdot}$ (so $A^n x \to 0$ in $\mathcal{X}$ for some $x \in \mathcal{X}$ implies that $x = 0$). Indeed, if A is such an operator, set

$C = (I - A^*A)^{1/2}$ considered as an operator from $\mathcal{X}$ into $\mathcal{Y} := \overline{\text{Ran}}(I - A^*A)^{1/2}$ (the closure of the range of $(I - A^*A)^{1/2}$). Such an A is not strongly stable by the definition of the class $C_{1\cdot}$, the definition of C makes the pair (C, A) isometric, and the condition that A is c.n.i. implies that (C, A) is observable.

To construct such an operator A, let θ be a Schur-class outer function such that $\log(1 - |\theta|^2)$ is not integrable (with respect to arc-length Lebesgue measure) over the unit circle $\mathbb{T}$. Furthermore, let $\mathbb{K}(\theta)$ be the associated Sz.-Nagy–Foiaş model space

$$\mathbb{K}(\theta) = \begin{bmatrix} H^2(\mathbb{D}) \\ L^2(\mathbb{T}) \end{bmatrix} \ominus \begin{bmatrix} \theta(\lambda) \\ (1 - |\theta(\zeta)|^2)^{1/2} \end{bmatrix} H^2(\mathbb{D}), \quad \text{where } \lambda \in \mathbb{D} \text{ and } \zeta \in \mathbb{T}$$

and let $S(\theta)$ be the Sz.-Nagy–Foiaş model operator

$$T = P_{\mathbb{K}(\theta)} \begin{bmatrix} M_\lambda & 0 \\ 0 & M_\zeta \end{bmatrix}\Bigg|_{\mathbb{K}(\theta)},$$

where M_λ and M_ζ are the operators of multiplication by λ and by ζ, respectively. Now we let $A := S(\theta)^*$ and note that A is in the class $C_{1\cdot}$ by Proposition 3.5 in Sz.-Nagy and Foiaş (1970) (since θ is outer) and A is c.n.i. by Theorem 5 in Ball and Kriete (1987) (since the nonlog-integrability property of $1-|\theta|^2$ implies that there is no H^∞-function $a(z)$ for which $|a(\zeta)|^2 \le 1 - |\theta(\zeta)|^2$ for $\zeta \in \mathbb{T}$). This completes the construction.

Let us say that *the pair* $(C, \mathbf{A})$ *is similar to the pair* $(\widetilde{C}, \widetilde{\mathbf{A}})$ if there is an invertible operator S on $\mathcal{X}$ so that

$$\widetilde{C} = CS^{-1}, \qquad \widetilde{A}_j = SA_jS^{-1} \quad \text{for } j = 1, \ldots, d.$$

Then we have the following characterization of pairs $(C, \mathbf{A})$ which are similar to a contractive or to an isometric pair.

Proposition 2.5

(1) *The pair* $(C, \mathbf{A})$ *is similar to a contractive pair* $(\widetilde{C}, \widetilde{\mathbf{A}})$ *if and only if there exists a bounded, strictly positive-definite solution* H *to the Stein inequality* (2.14).
(2) *The pair* $(C, \mathbf{A})$ *is similar to an isometric pair if and only if there exists a bounded, strictly positive-definite solution* H *of the Stein equation* (2.15).

Proof Suppose that H is a strictly positive-definite solution of (2.14). Factor H as $H = S^*S$ with S invertible and set

$$\widetilde{C} = CS^{-1}, \qquad \widetilde{A}_k = SA_kS^{-1} \quad \text{for } k = 1, \ldots, d. \tag{2.21}$$

Multiplying (2.14) on the left by S^{*-1} and on the right by S^{-1} then leads us to

$$I - \widetilde{A}_1^*\widetilde{A}_1 - \cdots - \widetilde{A}_d^*\widetilde{A}_d \ge \widetilde{C}^*\widetilde{C},$$

i.e., $(\widetilde{C}, \widetilde{\mathbf{A}})$ is a contractive pair which is similar to the original pair $(C, \mathbf{A})$. Conversely, if $(\widetilde{C}, \widetilde{\mathbf{A}})$ given by (2.21) is contractive, then $H = S^*S$ is bounded and positive-definite and satisfies the Stein inequality (2.14). This verifies the first statement of the Proposition. The second statement follows in a similar way. □

As a consequence of the observations in Proposition 2.5, Proposition 2.3 can be formulated more generally as follows.

Proposition 2.6

(1) *If the pair $(C, \mathbf{A})$ is such that the Stein inequality (2.14) has a strictly positive-definite solution H, then $(C, \mathbf{A})$ is output-stable. Moreover, the observability gramian $\mathcal{G}_{C,\mathbf{A}}$ is the unique positive semidefinite solution of the Stein equation (2.15) if and only if $\mathbf{A}$ is strongly stable.*

(2) *If the pair $(C, \mathbf{A})$ is such that the Stein equation (2.15) has a strictly positive-definite solution H, then $(C, \mathbf{A})$ is output-stable and the observability gramian $\mathcal{G}_{C,\mathbf{A}}$ is the unique positive semidefinite solution of the Stein equation (2.15) if and only if $\mathbf{A}$ is strongly stable. In this case $(C, \mathbf{A})$ is moreover exactly observable.*

The last part of Proposition 2.6 has a converse.

Proposition 2.7 *Suppose that the pair $(C, \mathbf{A})$ is output-stable and exactly observable. Then $\mathbf{A}$ is strongly stable, i.e., (2.11) holds.*

Proof If $(C, \mathbf{A})$ is output-stable and exactly observable, then the observability gramian $\mathcal{G}_{C,\mathbf{A}}$ is a strictly positive-definite solution of the Stein equation (2.15). Hence (2.19) holds with $H = \mathcal{G}_{C,\mathbf{A}}$:

$$\langle \mathcal{G}_{C,\mathbf{A}} x, x\rangle = \sum_{v \in \mathcal{F}_d\colon |v| \le N} \langle \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v x, x\rangle + \sum_{v \in \mathcal{F}_d\colon |v| = N+1} \langle \mathcal{G}_{C,\mathbf{A}} \mathbf{A}^v x, \mathbf{A}^v x\rangle. \tag{2.22}$$

From the infinite-series representation (2.10) for $\mathcal{G}_{C,\mathbf{A}}$, taking limits in (2.22) gives

$$\lim_{N\to\infty} \sum_{v \in \mathcal{F}_d\colon |v| = N+1} \langle \mathcal{G}_{C,\mathbf{A}} \mathbf{A}^v x, \mathbf{A}^v x\rangle = 0. \tag{2.23}$$

The strict positive-definiteness of $\mathcal{G}_{C,\mathbf{A}}$ tells us that there is an $\varepsilon > 0$ so that

$$\varepsilon \|x\|^2 \le \langle \mathcal{G}_{C,\mathbf{A}} x, x\rangle \quad \text{for all } \ x \in \mathcal{X}. \tag{2.24}$$

In particular, from (2.24) with $\mathbf{A}^v x$ in place of x combined with (2.23) we get

$$\varepsilon \sum_{v \in \mathcal{F}_d\colon |v| = N+1} \|\mathbf{A}^v x\|^2 \le \sum_{v \in \mathcal{F}_d\colon |v| = N+1} \langle \mathcal{G}_{C,\mathbf{A}} \mathbf{A}^v x, \mathbf{A}^v x\rangle \to 0$$

for all $x \in \mathcal{X}$, and we conclude that $\mathbf{A}$ is strongly stable as asserted. □

2.2 Observability-operator range spaces and reproducing kernel Hilbert spaces: the noncommutative-variable case

To develop the noncommutative analogue of Theorem 1.2, we first introduce the right noncommutative shift operators $S_1^R, \ldots, S_d^R$ on $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ as follows:

$$S_j^R\colon \sum_{v \in \mathcal{F}_d} f_v z^v \mapsto \sum_{v \in \mathcal{F}_d} f_v z^v z_j = \sum_{v \in \mathcal{F}_d} f_v z^{vj} \quad (j = 1, \ldots, d). \tag{2.25}$$

It is readily seen that their adjoints (backward shifts) are given by

$$(S_j^R)^*\colon \sum_{v \in \mathcal{F}_d} f_v z^v \mapsto \sum_{v \in \mathcal{F}_d} f_{vj} z^v \quad (j = 1, \ldots, d). \tag{2.26}$$

Their left counterparts $S_1^L, \ldots, S_d^L$, also on $H^2_{\mathcal{Y}}(\mathcal{F}_d)$, are given by

$$S_j^L\colon \sum_{v \in \mathcal{F}_d} f_v z^v \mapsto \sum_{v \in \mathcal{F}_d} f_v z_j z^v = \sum_{v \in \mathcal{F}_d} f_v z^{jv} \tag{2.27}$$

with adjoints given by

$$(S_j^L)^*\colon \sum_{v\in\mathcal{F}_d} f_v z^v \mapsto \sum_{v\in\mathcal{F}_d} f_{jv} z^v. \tag{2.28}$$

Let τ denote the unitary involution on $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ given by

$$\tau\colon \sum_{v\in\mathcal{F}_d} f_v z^v \to \sum_{v\in\mathcal{F}_d} f_{v^\top} z^v. \tag{2.29}$$

In addition to the unitary property $\tau^* = \tau^{-1}$ of τ, note also that τ intertwines the left shifts with the right shifts:

$$(S_j^R)^*\tau = \tau(S_j^L)^*, \quad S_j^R\tau = \tau S_j^L \quad \text{for} \quad j = 1, \ldots, d. \tag{2.30}$$

Then we have the following Fock-space analogue of Theorem 1.2.

Theorem 2.8 *Suppose that* $(C, \mathbf{A})$ *is an output-stable pair. Then:*

(1) *The intertwining relation*

$$(S_j^R)^*\widehat{\mathcal{O}}_{C,\mathbf{A}}x = \widehat{\mathcal{O}}_{C,\mathbf{A}}A_j x \quad (x \in \mathcal{X}) \tag{2.31}$$

holds for every backward-shift operator $(S_j^R)^*$ *defined in* (2.26) *and hence* $\operatorname{Ran}\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *is* $(S_j^R)^*$*-invariant for* $j = 1, \ldots, d$.

(2) *Let* $H \geq 0$ *be a solution of the Stein inequality* (2.14) *and let* $\mathcal{X}'$ *be the completion of* $\mathcal{X}$ *with H-inner product* $\|[x]\|^2_{\mathcal{X}'} = \langle Hx, x\rangle_{\mathcal{X}}$. *Then* A_j *and* C *extend to define bounded operators* $A'_j\colon \mathcal{X}' \to \mathcal{X}'$ *for* $j = 1, \ldots, d$ *and* $C'\colon \mathcal{X}' \to \mathcal{Y}$ *and the observability operator* $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *extends to define a contraction operator* $\widehat{\mathcal{O}}_{C',\mathbf{A}'}$ *from* $\mathcal{X}'$ *into* $H^2_{\mathcal{Y}}(\mathcal{F}_d)$. *Moreover,* $\widehat{\mathcal{O}}_{C',\mathbf{A}'}\colon \mathcal{X}' \to H^2_{\mathcal{Y}}(\mathcal{F}_d)$ *is an isometry if and only if* H *satisfies the Stein equation* (2.15) *and* $\mathbf{A}' = (A'_1, \ldots, A'_d)$ *is strongly stable, i.e.,*

$$\sum_{v\in\mathcal{F}_d\colon |v|=N} \langle HA^v x, A^v x\rangle \to 0 \quad \textit{for all} \;\; x \in \mathcal{X}. \tag{2.32}$$

(3) *If* $H \geq 0$ *is a solution of the Stein inequality* (2.14) *and the linear manifold* $\mathcal{M} := \operatorname{Ran}\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *is given the lifted norm*

$$\|\widehat{\mathcal{O}}_{C,\mathbf{A}}x\|^2_{\mathcal{M}} = \inf_{y\in\mathcal{X}\colon \widehat{\mathcal{O}}_{C,\mathbf{A}}y = \widehat{\mathcal{O}}_{C,\mathbf{A}}x} \langle Hy, y\rangle_{\mathcal{X}}, \tag{2.33}$$

then

(*a*) $\mathcal{M}$ *can be completed to* $\mathcal{M}' = \operatorname{Ran}\widehat{\mathcal{O}}_{C',\mathbf{A}'}$ *(with* $(C', \mathbf{A}')$ *as in #2 above) with contractive inclusion in* $H^2_{\mathcal{Y}}(\mathcal{F}_d)$*:*

$$\|f\|^2_{H^2_{\mathcal{Y}}(\mathcal{F}_d)} \leq \|f\|^2_{\mathcal{M}'} \quad \textit{for all} \quad f \in \mathcal{M}'.$$

Furthermore, $\mathcal{M}'$ *is isometrically equal to the formal noncommutative reproducing kernel Hilbert space with reproducing kernel* $K_{C,\mathbf{A};H}$ *given by* (1.16).

(*b*) *The following difference-quotient inequality is valid*

$$\sum_{j=1}^{d} \|(S_j^R)^* f\|^2_{\mathcal{H}(K_{C,\mathbf{A};H})} \leq \|f\|^2_{\mathcal{H}(K_{C,\mathbf{A};H})} - \|f_\emptyset\|^2_{\mathcal{Y}} \tag{2.34}$$

for every $f \in \mathcal{M}' = \mathcal{H}(K_{C,\mathbf{A};H})$ *with equality holding in* (2.34) *if and only if* (2.15) *holds.*

(4) *Conversely, if $\mathcal{M}$ is a Hilbert space included in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ which is invariant under $(S^R_j)^*$ for $j = 1, \ldots, d$ such that the difference-quotient inequality*

$$\sum_{j=1}^{d} \|(S^R_j)^* f\|^2_{\mathcal{M}} \leq \|f\|^2_{\mathcal{M}} - \|f_\emptyset\|^2_{\mathcal{Y}} \tag{2.35}$$

holds for every $f \in \mathcal{M}$, then $\mathcal{M}$ is contractively included in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ and there exists a contractive pair $(C, \mathbf{A})$ (so $H = I$ positive definite solution of the Stein inequality (2.14)) such that

$$\mathcal{M} = \mathcal{H}(K_{C,\mathbf{A};I}) = \operatorname{Ran} \widehat{\mathcal{O}}_{C,\mathbf{A}}$$

isometrically. In case (2.35) holds with equality, then $(C, \mathbf{A})$ can be taken to be an isometric pair. An explicit $(C, \mathbf{A})$ meeting these conditions is given as follows. Take $\mathcal{X}$ to be the Hilbert space $\mathcal{X} = \tau(\mathcal{M})$ (where τ is the involution given by (2.29)) with $\|\tau(f)\|_{\mathcal{X}} = \|f\|_{\mathcal{M}}$ and define $C\colon \mathcal{X} \to \mathcal{Y}$ and $\mathbf{A} = (A_1, \ldots, A_d)$ on $\mathcal{X}$ by

$$A_j = (S^L_j)^*|_{\mathcal{X}} \text{ for } j = 1, \ldots, d; \qquad C = E|_{\mathcal{X}}\colon \mathcal{X} \to \mathcal{Y}, \qquad H = I_{\mathcal{X}}, \tag{2.36}$$

where $E\colon H^2_{\mathcal{Y}}(\mathcal{F}_d) \to \mathcal{Y}$ is given by

$$E\colon \sum_{v \in \mathcal{F}_d} f_v z^v \mapsto f_\emptyset. \tag{2.37}$$

Proof of (1) Applying $(S^R_j)^*$ to a typical element from $\operatorname{Ran} \widehat{\mathcal{O}}_{C,\mathbf{A}}$ (with notation as in (2.7)), we get

$$\begin{aligned}(S^R_j)^*(C(I - Z(z)A)^{-1}x) &= (S^R_j)^*\Big(\sum_{v \in \mathcal{F}_d} (C\mathbf{A}^v x) z^v\Big) \\ &= \sum_{v \in \mathcal{F}_d} (C\mathbf{A}^{vj} x) z^v \\ &= C(I - Z(z)A)^{-1} A_j x \in \operatorname{Ran} \widehat{\mathcal{O}}_{C,\mathbf{A}}.\end{aligned} \tag{2.38}$$

The latter equality shows that $\operatorname{Ran} \widehat{\mathcal{O}}_{C,\mathbf{A}}$ is invariant under $(S^R_j)^*$ for all $j = 1, \ldots, d$ (backward-shift-invariant) and (2.31) follows. □

Proof of (2) The Stein inequality (2.14) amounts to the statement that $(C, \mathbf{A})$ is contractive and well-defined on the dense subset $[\mathcal{X}]$ of $\mathcal{X}'$ (where $[x]$ is the equivalence class containing x) and hence extends to a contractive pair $(C', \mathbf{A}')$ on all of $\mathcal{X}'$ and moreover the inequality (2.17) holds for all $N = 1, 2, \ldots$ From this we see that $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ is contractive from $\mathcal{X}$ with the H-inner product to $H^2_{\mathcal{Y}}(\mathcal{F}_d)$, and hence also $\widehat{\mathcal{O}}_{C',\mathbf{A}'}$ is contractive from $\mathcal{X}'$ to $H^2_{\mathcal{Y}}(\mathcal{F}_d)$. The inequality (2.17) is actually a chain of inequalities $W_N \geq W_{N+1}$ for $N = 1, 2, \ldots$, where

$$W_N = \sum_{v \in \mathcal{F}_d\colon |v| \leq N-1} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v + \sum_{v \in \mathcal{F}_d\colon |v| = N} \mathbf{A}^{*v^\top} H \mathbf{A}^v.$$

Note that

$$\text{s-lim}_{N \to \infty} W_N = \mathcal{G}_{C,\mathbf{A}} + \Delta_{H,\mathbf{A}},$$

where

$$\Delta_{H,\mathbf{A}} = \text{s-lim}_{N \to \infty} \sum_{|v|=N} \mathbf{A}^{*v^\top} H \mathbf{A}.$$

In particular it follows from (2.16) that

$$H \geq C^*C + \sum_{j=1}^{d} A_j^* H A_j \geq W_N \quad \text{for all } \; N \geq 2$$

and hence, by taking the strong limit on the right hand side, we get

$$H \geq C^*C + \sum_{j=1}^{d} A_j^* H A_j \geq \mathcal{G}_{C,\mathbf{A}} + \Delta_{H,\mathbf{A}}. \tag{2.39}$$

By definition, $\widehat{\mathcal{O}}_{C',\mathbf{A}'} \colon \mathcal{X}' \to H^2_{\mathcal{Y}}(\mathcal{F}_d)$ being an isometry means that $\mathcal{G}_{C,\mathbf{A}} = H$ in which case (2.39) becomes

$$H \geq C^*C + \sum_{j=1}^{d} A_j^* H A_j \geq H + \Delta_{H,\mathbf{A}}, \tag{2.40}$$

which in turn forces $\Delta_{H,\mathbf{A}} = 0$ and equalities throughout (2.40). The condition $\Delta_{H,\mathbf{A}} = 0$ just means that $\mathbf{A}'$ is strongly stable. From equality holding in (2.40) we see that the Stein inequality (2.14) holds with equality, i.e., the Stein equation (2.15) holds. Conversely, by reversing the steps of the argument, we see that $\mathbf{A}'$ being strongly stable and the Stein equality holding leads to $\mathcal{G}_{C,\mathbf{A}} = H$, i.e., to $\widehat{\mathcal{O}}_{C',\mathbf{A}'}$ being an isometry from $\mathcal{X}'$ into $H^2_{\mathcal{Y}}(\mathcal{F}_d)$. □

Proof of (3a) Statement (3a) follows from general principles laid out in Ball and Vinnikov (2003).

Proof of (3b) For f of the form $f(z) = C(I - Z(z)A)^{-1}x$, we have

$$\|f\|^2_{\mathcal{H}(K_{C,\mathbf{A};H})} = \langle Hx, x\rangle.$$

We see that then

$$\|S_j^* f\|^2_{\mathcal{H}(K_{C,\mathbf{A};H})} = \langle HA_j x, A_j x\rangle \text{ for } \; j = 1, \ldots, d \text{ (from (2.38))}, \qquad f_\emptyset = Cx.$$

With these substitutions, we see that (2.34) is equivalent to

$$\sum_{j=1}^{d} \langle HA_j x, A_j x\rangle_{\mathcal{X}} \leq \langle Hx, x\rangle_{\mathcal{X}} - \|Cx\|^2_{\mathcal{Y}}$$

or, in operator-theoretic form,

$$A_1^* H A_1 + \cdots + A_d^* H A_d \leq H - C^*C \tag{2.41}$$

with equality in (2.34) equivalent to equality in (2.41). This completes the verification of part (3b) of Theorem 2.8. □

Before commencing the proof of part (4) of Theorem 2.8, we collect some useful facts concerning $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ itself.

Proposition 2.9 *Let* $\mathbf{S} = (S_1, \ldots, S_d)$ *denote either the right shift* $\mathbf{S}^R$ *or the left shift* $\mathbf{S}^L$

$$\mathbf{S}^R = (S_1^R, \ldots, S_d^R), \qquad \mathbf{S}^L = (S_1^L, \ldots, S_d^L)$$

defined as in (2.25) *and* (2.27) *and let the operator* E: $H^2_{\mathcal{Y}}(\mathcal{F}_d), \to \mathcal{Y}$ *be defined as in* (2.37). *Then:*

(1) *The operator-tuple* $\mathbf{S}^* = (S_1^*, \ldots, S_d^*)$ *is strongly stable, i.e.,*

$$\lim_{N\to\infty} \sum_{v\in\mathcal{F}_d\colon |v|=N} \|\mathbf{S}^{*v}f\|^2 = 0 \quad \textit{for each} \quad f \in H^2_{\mathcal{Y}}(\mathcal{F}_d). \tag{2.42}$$

(2) *The operator*

$$\begin{bmatrix} S_1^* \\ \vdots \\ S_d^* \\ E \end{bmatrix} \colon H^2_{\mathcal{Y}}(\mathcal{F}_d) \to (H^2_{\mathcal{Y}}(\mathcal{F}_d))^d \oplus \mathcal{Y}$$

is unitary, i.e.,

$$EE^* = I_{\mathcal{Y}}, \quad ES_i = 0, \quad S_j^*S_i = \delta_{ij}I \quad \textit{for} \quad i,j = 1,\ldots,d \tag{2.43}$$

(where δ_{ij} *stands for the Kronecker symbol), and*

$$I - S_1S_1^* - \cdots - S_dS_d^* = E^*E. \tag{2.44}$$

(3) $X = I$ *is the unique solution of the Stein equation*

$$X - S_1XS_1^* - \cdots - S_dXS_d^* = E^*E. \tag{2.45}$$

(4) *For every* $f \in H^2_{\mathcal{Y}}(\mathcal{F}_d)$*,*

$$f(z) - f_\emptyset = \sum_{j=1}^d (S_jS_j^*f)(z) = \begin{cases} \sum_{j=1}^d ((S_j^R)^*f)(z)\cdot z_j, & \textit{if } \mathbf{S} = \mathbf{S}^R, \\ \sum_{j=1}^d z_j \cdot ((S_j^L)^*f)(z), & \textit{if } \mathbf{S} = \mathbf{S}^L. \end{cases} \tag{2.46}$$

(5) *The observability operator* $\widehat{\mathcal{O}}_{E,\mathbf{S}^{R*}}$ *is equal to the operator* τ *defined in* (2.29) *and hence is unitary.*

Proof If $f(z) = \sum_{v\in\mathcal{F}_d} f_v z^v$, then

$$(\mathbf{S}^{R*v'}f)(z) = \sum_{v\in\mathcal{F}_d} f_{vv'^\top} z^v, \qquad (\mathbf{S}^{L*v'}f)(z) = \sum_{v\in\mathcal{F}_d} f_{v'^\top v} z^v$$

and hence, in either the left or the right case, we have

$$E\mathbf{S}^{*v}f = f_{v^\top}. \tag{2.47}$$

Therefore,

$$\sum_{v\in\mathcal{F}_d\colon |v|=N} \|\mathbf{S}^{*v}f\|^2 = \sum_{v\in\mathcal{F}_d\colon |v|\geq N} \|f_v\|^2 \to 0 \quad \text{as} \quad N\to\infty,$$

and (2.42) follows. Equalities (2.43) and (2.44) follow from (2.25)–(2.28), (2.37) and the fact that E^* is the inclusion map of $\mathcal{Y}$ into $H^2_{\mathcal{Y}}(\mathcal{F}_d)$. Applying the operator identity (2.44) to an $f \in H^2_{\mathcal{Y}}(\mathcal{F}_d)$, we get (2.46). Finally, from (2.47) we see that, for both the left and the right case,

$$\widehat{\mathcal{O}}_{E,\mathbf{S}^*}f = \sum_{v\in\mathcal{F}_d} (E\mathbf{S}^{*v}f)z^v = \sum_{v\in\mathcal{F}_d} f_{v^\top}z^v = \tau f$$

for all $f \in H^2_{\mathcal{Y}}(\mathcal{F}_d)$. That $X = I$ is the unique solution of the Stein equation (2.45) is now a consequence of (2.42) combined with the last part of Theorem 2.2. □

Proof of (4) in Theorem 2.8 Suppose that $\mathcal{M}$ is a Hilbert space contractively included in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ which is invariant under $(S^R_j)^*$ for each $j = 1, \dots, d$ such that the difference-quotient inequality (2.34) holds. Set $\mathcal{X} = \tau(\mathcal{M})$ with norm inherited from $\mathcal{M}$. From the intertwining relations (2.30) we see that $\mathcal{X}$ is invariant under the left backward shifts $(S^L_1)^*, \dots, (S^L_d)^*$. Define operators A_j: $\mathcal{X} \to \mathcal{X}$ for $j = 1, \dots, d$ and C: $\mathcal{X} \to \mathcal{Y}$ by (2.36). From the difference-quotient inequality (2.35) together with the definition of the $\mathcal{X}$-norm and the intertwining relations (2.30), we have

$$\begin{aligned}\sum_{j=1}^{d} \|(S^L_j)^*(\tau f)\|^2_{\mathcal{X}} &= \sum_{j=1}^{d} \|\tau(S^R_j)^* f)\|^2_{\mathcal{X}} = \sum_{j=1}^{d} \|(S^R_j)^* f\|^2_{\mathcal{M}} \\ &\leq \|f\|^2_{\mathcal{M}} - \|f_\emptyset\|^2_{\mathcal{Y}} = \|\tau(f)\|^2_{\mathcal{X}} - \|(\tau f)_\emptyset\|^2_{\mathcal{Y}}\end{aligned}$$

and hence $H = I_{\mathcal{X}}$ satisfies the Stein inequality (2.14). From Proposition 2.9 we see that

$$\widehat{\mathcal{O}}_{C,\mathbf{A}} = \widehat{\mathcal{O}}_{E,\mathbf{S}^{L*}}|_{\mathcal{X}} = \tau|_{\mathcal{X}=\tau(\mathcal{M})}$$

and hence $\widehat{\mathcal{O}}_{C,\mathbf{A}}\tau|_{\mathcal{M}} = I_{\mathcal{M}}$. Therefore, for each $f \in \mathcal{M}$ we have

$$\|f\|_{\mathcal{H}(K_{C,\mathbf{A};I})} = \|\widehat{\mathcal{O}}_{C,\mathbf{A}}\tau f\|_{\mathcal{H}(K_{C,\mathbf{A}})} = \|\tau f\|_{\mathcal{X}} = \|f\|_{\mathcal{M}}$$

and thus $\mathcal{M} = \mathcal{H}(K_{C,\mathbf{A}})$ isometrically. It then follows from part (3a) of the theorem that in fact $\mathcal{M}$ is contractively included in $H_{\mathcal{Y}}(\mathcal{F}_d)$. □

As explained by part (4) of Theorem 2.8, for purposes of study of contractively included, backward-shift-invariant subspaces of $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ which satisfy the difference-quotient-inequality (2.34), without loss of generality we may suppose at the start that we are working with $\mathcal{X}'$ as the original state space $\mathcal{X}$ and with the solution H of the Stein inequality (2.14) to be normalized to $H = I_{\mathcal{X}}$. Then certain simplifications occur in parts (1)-(4) of Theorem 2.8 as explained in the next result.

Theorem 2.10 *Suppose that* $(C, \mathbf{A})$ *is a contractive pair with state space* $\mathcal{X}$ *and output space* $\mathcal{Y}$. *Then:*

(1) $(C, \mathbf{A})$ *is output-stable and the intertwining relation* (2.31) *holds. Hence* Ran $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *is invariant under the backward shifts* $(S^R_j)^*$ *for* $j = 1, \dots, d$.
(2) *The observability operator* $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *is a contraction from* $\mathcal{X}$ *into* $H^2_{\mathcal{Y}}(\mathcal{F}_d)$. *Moreover* $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *is isometric if and only if* $(C, \mathbf{A})$ *is an isometric pair and* $\mathbf{A}$ *is strongly stable.*
(3) *If the linear manifold* $\mathcal{M} :=$ Ran $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *is given the lifted norm*

$$\|\widehat{\mathcal{O}}_{C,\mathbf{A}}x\|_{\mathcal{M}} = \|Qx\|_{\mathcal{X}}, \tag{2.48}$$

where Q *is the orthogonal projection of* $\mathcal{X}$ *onto* $(\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^\perp$, *then* $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ *is a coisometry of* $\mathcal{X}$ *onto* $\mathcal{M}$. *Moreover,* $\mathcal{M}$ *is contained contractively in* $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ *and is isometrically equal to the formal noncommutative reproducing kernel Hilbert space* $\mathcal{H}(K_{C,\mathbf{A}})$ *with reproducing kernel* $K_{C,\mathbf{A}}(z, w)$ *given by*

$$K_{C,\mathbf{A}}(z, w) = C(I - Z(z)A)^{-1}(I - A^* Z(w)^*)^{-1} C^*.$$

(4) *If $\mathcal{O}_{C,\mathbf{A}}$ is given the lifted norm $\|\cdot\|_{\mathcal{H}(K_{C,\mathbf{A}})}$ as in (2.48), then the difference-quotient inequality*

$$\sum_{j=1}^{d} \|(S_j^R)^* f\|^2_{\mathcal{H}(K_{C,\mathbf{A}})} \leq \|f\|^2_{\mathcal{H}(K_{C,\mathbf{A}})} - \|f_\emptyset\|^2_{\mathcal{Y}} \tag{2.49}$$

holds for all $f \in \mathcal{H}(K_{C,\mathbf{A}})$. Moreover, (2.49) holds with equality if and only the orthogonal projection Q of $\mathcal{X}$ onto $(\operatorname{Ker} \mathcal{O}_{C,\mathbf{A}})^\perp$ satisfies the Stein equation

$$Q - \sum_{j=1}^{d} A_j^* Q A_j = C^* C. \tag{2.50}$$

In particular, if $(C, \mathbf{A})$ is observable, then (2.49) holds with equality if and only if $(C, \mathbf{A})$ is an isometric pair.

Proof Statements (1)–(3) and all but the last part of statement (4) are direct specializations to the case $H = I_{\mathcal{X}}$ of the corresponding results in Theorem 2.8. It remains only to analyze the conditions for equality in (2.49).

From the intertwining relation (2.31), we see that equality in (2.49) for a generic element $f = \widehat{\mathcal{O}}_{C,\mathbf{A}} x \in \mathcal{H}(K_{C,\mathbf{A}})$ means that

$$\sum_{j=1}^{d} \|\widehat{\mathcal{O}}_{C,\mathbf{A}} A_j x\|^2_{\mathcal{H}(K_{C,\mathbf{A}})} = \|\widehat{\mathcal{O}}_{C,\mathbf{A}} x\|^2_{\mathcal{H}(K_{C,\mathbf{A}})} - \|Cx\|^2_{\mathcal{Y}}$$

for all $x \in \mathcal{X}$. Using the definition (2.48) of the $\mathcal{H}(K_{C,\mathbf{A}})$-norm, we rewrite this last equality as

$$\sum_{j=1}^{d} \|Q A_j x\|^2_{\mathcal{X}} = \|Qx\|^2_{\mathcal{X}} - \|Cx\|^2_{\mathcal{Y}}.$$

This holding for all $x \in \mathcal{X}$ is finally equivalent to the Stein equation (2.50). □

Remark 2.11 In Theorems 2.8 and 2.10, we could equally well have interchanged the roles of left versus right. For a given output pair $(C, \mathbf{A})$, define the associated *left observability operator* $\mathcal{O}^L_{C,\mathbf{A}} : \mathcal{X} \to H^2_{\mathcal{Y}}(\mathcal{F}_d)$ by

$$\mathcal{O}^L_{C,\mathbf{A}} x = \sum_{v \in \mathcal{F}_d} C \mathbf{A}^{v^\top} z^v.$$

Then the linear manifold $\operatorname{Ran} \mathcal{O}^L_{C,\mathbf{A}}$ is invariant under the left backward shifts $((S_1^L)^*, \ldots, (S_d^L)^*)$ as verified by the intertwining relation

$$(S_j^L)^* \mathcal{O}^L_{C,\mathbf{A}} = \mathcal{O}^L_{C,\mathbf{A}} A_j.$$

We leave the precise statements and proofs to the interested reader.

The characterization (2.50) of the difference-quotient inequality holding with equality for a space $\mathcal{H}(K_{C,\mathbf{A}})$ in Theorem 2.10 can be made more explicit as follows.

Proposition 2.12 *Suppose that $(C, \mathbf{A})$ is a contractive pair as in Theorem 2.10 and let Q be the orthogonal projection onto $(\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^\perp$. Then Q satisfies the Stein inequality*

$$Q - A_1^* Q A_1 - \cdots A_d^* Q A_d \geq C^* C \tag{2.51}$$

and we have the inequalities

$$\mathcal{G}_{C,\mathbf{A}} \leq Q \leq I_{\mathcal{X}}. \tag{2.52}$$

If we write C, A_j, Q in 2×2-block matrix form with respect to the decomposition $\mathcal{X} = \operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}} \oplus (\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^{\perp}$ as

$$C = \begin{bmatrix} 0 & C^0 \end{bmatrix}, \qquad A_j = \begin{bmatrix} A_{j1} & A_{j2} \\ 0 & A_j^0 \end{bmatrix}, \qquad Q = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \tag{2.53}$$

for $j = 1, \ldots, d$, then Q satisfies the Stein equation (2.50) if and only if the pair $(C^0, \mathbf{A}^0)$ is an isometric pair, in which case we also have that $A_{j2} = 0$ (so $(\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^{\perp}$ is invariant for A_j) for $j = 1, \ldots, d$.

Proof First note that $\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}}$ is invariant for each A_j and that $\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}} \subset \operatorname{Ker} C$. Therefore, the matrix decompositions of C, A_j, Q with respect to the decomposition $\mathcal{X} = \operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}} \oplus (\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^{\perp}$ have the form as given in (2.53). Next note that the contractive property of the pair $(C, \mathbf{A})$ means that

$$\begin{bmatrix} 0 & 0 \\ 0 & C^{0*}C^0 \end{bmatrix} + \sum_{j=1}^{d} \begin{bmatrix} A_{j1}^* A_{j1} & A_{j1}^* A_{j2} \\ A_{j2}^* A_{j1} & A_{j2}^* A_{j2} + A_j^{0*} A_j^{0*} \end{bmatrix} \leq \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{2.54}$$

On the other hand, the Stein inequality (2.51) works out to be

$$\begin{bmatrix} 0 & 0 \\ 0 & C^{0*}C^0 \end{bmatrix} + \sum_{j=1}^{d} \begin{bmatrix} A_{j1}^* A_{j1} & A_{j1}^* A_{j2} \\ A_{j2}^* A_{j1} & A_j^{0*} A_j^{0*} \end{bmatrix} \leq \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{2.55}$$

As the left hand side of (2.55) is dominated by the left hand side of (2.54), it is clear that (2.55) follows from (2.54), and hence (2.51) holds as asserted. Since $\mathcal{G}_{C,\mathbf{A}}$ is the minimal positive semidefinite solution of the Stein inequality (2.14) (by part (2) of Theorem 2.2) and we now know that Q is one such solution, it follows that $\mathcal{G}_{C,\mathbf{A}} \leq Q$. As Q is an orthogonal projection on $\mathcal{X}$, we also have $Q \leq I_{\mathcal{X}}$ and (2.52) now follows.

From the $(2, 2)$ entry of (2.54), we read off

$$C^{0*}C^0 + \sum_{j=1}^{d} A_{j2}^* A_{j2} + \sum_{j=1}^{d} A_j^{0*} A_j^0 \leq I_{(\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^{\perp}}. \tag{2.56}$$

In particular

$$C^{0*}C^0 + \sum_{j=1}^{d} A_j^{0*} A_j^0 \leq I_{(\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^{\perp}},$$

i.e., Q satisfies (2.51). On the other hand, the validity of (2.50) reduces to

$$\begin{bmatrix} 0 & 0 \\ 0 & C^{0*}C^0 \end{bmatrix} + \sum_{j=1}^{d} \begin{bmatrix} 0 & 0 \\ 0 & A_j^{0*} A_j^0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$$

or simply to

$$C^{0*}C^0 + \sum_{j=1}^{d} A_j^{0*} A_j^0 = I_{(\operatorname{Ker} \widehat{\mathcal{O}}_{C,\mathbf{A}})^{\perp}}. \tag{2.57}$$

Thus, the validity of (2.50) is equivalent to $(C^0, \mathbf{A}^0)$ being an isometric pair, in which case we also have that $A_{j2} = 0$. □

Finally, we have the following uniqueness result.

Theorem 2.13 *Suppose that* $(C, \mathbf{A})$ *and* $(\widetilde{C}, \widetilde{\mathbf{A}})$ *are two output-stable, observable pairs realizing the same positive kernel*

$$\begin{aligned} K_{C,\mathbf{A}}(z,w) &:= C(I - Z(z)A)^{-1}(I - A^*Z(w)^*)^{-1}C^* \\ &= \widetilde{C}(I - Z(z)\widetilde{A})^{-1}(I - \widetilde{A}^*Z(w)^*)^{-1}\widetilde{C}^* =: K_{\widetilde{C},\widetilde{\mathbf{A}}}(z,w). \end{aligned} \tag{2.58}$$

Then $(C, \mathbf{A})$ *and* $(\widetilde{C}, \widetilde{\mathbf{A}})$ *are* unitarily equivalent, *i.e., there is a unitary operator* $U\colon \mathcal{X} \to \widetilde{\mathcal{X}}$ *such that*

$$C = \widetilde{C}U \quad \textit{and} \quad A_j = U^{-1}\widetilde{A}_jU \quad \textit{for } j = 1, \ldots, d.$$

Proof For any two words $\alpha, \beta \in \mathcal{F}_d$, equating coefficients of $z^\alpha w^{\beta^\top}$ in (2.58) gives

$$C\mathbf{A}^\alpha\mathbf{A}^{*\beta}C^* = \widetilde{C}\widetilde{\mathbf{A}}^\alpha\widetilde{\mathbf{A}}^{*\beta}\widetilde{C}^*.$$

Hence the operator U defined by

$$U\colon \mathbf{A}^{*\beta}C^*y \mapsto \widetilde{\mathbf{A}}^{*\beta}\widetilde{C}^*y \tag{2.59}$$

extends by linearity and continuity to define an isometry from

$$\mathcal{D}_U = \overline{\operatorname{span}}\{\mathbf{A}^{*\beta}C^*y\colon \beta \in \mathcal{F}_d,\ y \in \mathcal{Y}\}$$

onto

$$\mathcal{R}_U = \overline{\operatorname{span}}\{\widetilde{\mathbf{A}}^{*\beta}\widetilde{C}^*y\colon \beta \in \mathcal{F}_d,\ y \in \mathcal{Y}\}.$$

The observability assumption implies that $\mathcal{D}_U = \mathcal{X}$ and $\mathcal{R}_U = \widetilde{\mathcal{X}}$; hence $U\colon \mathcal{X} \to \widetilde{\mathcal{X}}$ is unitary. From (2.59) it is easily seen that

$$UC^* = \widetilde{C}^* \quad \text{and} \quad UA_j^* = \widetilde{A}_j^*U \quad \text{for } j = 1, \ldots, d.$$

Since U is unitary we then get

$$\widetilde{C}U = C \quad \text{and} \quad \widetilde{A}_jU = UA_j \quad \text{for } j = 1, \ldots, d$$

and we conclude that $(C, \mathbf{A})$ and $(\widetilde{C}, \widetilde{\mathbf{A}})$ are unitarily equivalent as desired.

□

2.3 Applications of observability operators: the noncommutative setting

As an application we give a proof of the Beurling-Lax theorem for the Fock-space setting originally given by Popescu (1989d). We shall in fact prove a more general version of the Beurling–Lax–Halmos theorem for contractively included (rather than isometrically included) subspaces of $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ due in the classical setting to de Branges (see de Branges & Rovnyak, 1966). Our proof is similar to that in Popescu (1989d) but highlights more explicitly the role of an associated observability operator. For this purpose we say that a formal power series $\theta(z) = \sum_{v\in\mathcal{F}_d}\theta_v z^v \in \mathcal{L}(\mathcal{U}, \mathcal{Y})\langle\langle z\rangle\rangle$ is a *contractive multiplier*, also written as θ is in the d-variable, noncommutative Schur-class $\mathcal{S}_{nc,d}(\mathcal{U}, \mathcal{Y})$, if the operator M_θ of multiplication by θ

$$M_\theta\colon f(z) \mapsto \theta(z) \cdot f(z)$$

defines a bounded linear operator from $H^2_{\mathcal{U}}(\mathcal{F}_d)$ to $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ with operator norm at most 1. Such a formal power series $\theta(z)$ is said to be *inner* if moreover the operator M_θ from $H^2_{\mathcal{U}}(\mathcal{F}_d)$ to $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ is an isometry.[1]

Theorem 2.14

(1) *A Hilbert space $\mathcal{M}$ is such that*
   (a) *$\mathcal{M}$ is contractively included in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$,*
   (b) *$\mathcal{M}$ is invariant under the right shift operators $S^R_1, \ldots, S^R_d$:*

$$S^R_j \mathcal{M} \subset \mathcal{M},$$

   (c) *the d-tuple*

$$\mathbf{S}^R_{\mathcal{M}} = (S^R_{\mathcal{M},1}, \ldots, S^R_{\mathcal{M},d}), \quad \textit{where} \;\; S^R_{\mathcal{M},j} := S^R_j|_{\mathcal{M}} \;\; \textit{for} \;\; j = 1, \ldots, d$$

   is a row contraction

$$S^R_{\mathcal{M},1}(S^R_{\mathcal{M},1})^* + \cdots + S^R_{\mathcal{M},d}(S^R_{\mathcal{M},d})^* \leq I_{\mathcal{M}}$$

   and
   (d) *$(\mathbf{S}^R_{\mathcal{M}})^*$ is strongly stable, i.e.,*

$$\lim_{n\to\infty} \sum_{v \in \mathcal{F}_d \colon |v|=n} \|(\mathbf{S}^R_{\mathcal{M}})^{*v} f\|^2_{\mathcal{M}} \to 0 \quad \textit{for all} \quad f \in \mathcal{M},$$

if and only if there is a coefficient Hilbert space $\mathcal{U}$ and a contractive multiplier $\theta \in \mathcal{S}_{nc,d}(\mathcal{U}, \mathcal{Y})$ so that

$$\mathcal{M} = \theta \cdot H^2_{\mathcal{U}}(\mathcal{F}_d)$$

with lifted norm

$$\|\theta \cdot f\|_{\mathcal{M}} = \|Qf\|_{H^2_{\mathcal{U}}(\mathcal{F}_d)}, \tag{2.60}$$

where Q is the orthogonal projection onto $(\mathrm{Ker}\ M_\theta)^\perp$.

(2) *The subspace $\mathcal{M}$ in part (1) above is isometrically included in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ if and only if the associated contractive multiplier θ is inner.*

Proof We first verify sufficiency in statement (1). Suppose that $\mathcal{M}$ has the form $\mathcal{M} = \theta \cdot H^2_{\mathcal{U}}(\mathcal{F}_d)$ for a contractive multiplier θ with $\mathcal{M}$-norm given by (2.60). From the fact that $\|M_\theta\| \leq 1$ it is easily verified that $\|\theta \cdot f\|_{H^2_{\mathcal{Y}}(\mathcal{F}_d)} \leq \|\theta \cdot f\|_{\mathcal{M}}$, i.e. (a) holds. From the intertwining property $S^R_j M_\theta = M_\theta S^R_j$ (note that S^R_j is multiplication by z_j on the right while M_θ is multiplication by θ on the left), property (b) follows.

If Q is the orthogonal projection onto $(\mathrm{Ker}\, M_\theta)^\perp \subset H^2_{\mathcal{U}}(\mathcal{F}_d)$, then the intertwining property $S^R_j M_\theta = M_\theta S^R_j$ implies that

$$QS^R_j = QS^R_j Q \quad \text{and} \quad (S^R_j)^* Q = Q(S^R_j)^* Q \quad \text{for} \quad j = 1, \ldots, d. \tag{2.61}$$

[1] We prefer to define *inner* to be isometric rather than partially isometric as in Popescu (1989d).

Thus

$$\left\| \begin{bmatrix} S^R_{\mathcal{M},1} & \cdots & S^R_{\mathcal{M},d} \end{bmatrix} \begin{bmatrix} \theta \cdot f_1 \\ \vdots \\ \theta \cdot f_d \end{bmatrix} \right\|^2_{\mathcal{M}} = \left\| \theta \begin{bmatrix} S^R_1 & \cdots & S^R_d \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_d \end{bmatrix} \right\|^2_{\mathcal{M}}$$

$$= \left\| Q \begin{bmatrix} S^R_1 & \cdots & S^R_d \end{bmatrix} \begin{bmatrix} Qf_1 \\ \vdots \\ Qf_d \end{bmatrix} \right\|^2_{H^2_{\mathcal{U}}(\mathcal{F}_d)}$$

$$\leq \left\| \begin{bmatrix} Qf_1 \\ \vdots \\ Qf_d \end{bmatrix} \right\|^2_{H^2_{\mathcal{U}}(\mathcal{F}_d)^d} = \left\| \begin{bmatrix} \theta \cdot f_1 \\ \vdots \\ \theta \cdot f_d \end{bmatrix} \right\|^2_{\mathcal{M}^d}$$

and property (c) follows. Finally, a short computation shows that

$$(S^R_{\mathcal{M},j})^* \colon\ \theta \cdot f \mapsto \theta \cdot S^{R*}_j Qf, \qquad (\mathbf{S}^R_{\mathcal{M}})^{*v} \colon\ \theta \cdot f \mapsto \theta \cdot \mathbf{S}^{R*v} Qf$$

and hence

$$\sum_{v \in \mathcal{F}_d \colon |v|=n} \|(\mathbf{S}^R_{\mathcal{M}})^{*v} \theta \cdot f\|^2_{\mathcal{M}} = \sum_{v \in \mathcal{F}_d \colon |v|=n} \|\mathbf{S}^{R*v} Qf\|^2_{H^2_{\mathcal{U}}(\mathcal{F}_d)} \to 0$$

as $n \to \infty$, and property (d) follows as well. Moreover, if θ is inner and $\mathcal{M} = \theta \cdot H^2_{\mathcal{U}}(\mathcal{F}_d)$ with the lifted norm (2.60), it is clear that $\mathcal{M}$ is contained in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ isometrically. This completes the proof of sufficiency in Theorem 2.14.

Suppose now that the Hilbert space $\mathcal{M}$ satisfies conditions (a), (b), (c), (d) in statement (1) of Theorem 2.14. Define a d-tuple of operators $\mathbf{A} = (A_1, \ldots, A_d)$ on $\mathcal{M}$ by

$$A_j = (S^R_{\mathcal{M},j})^* \quad \text{for} \quad j = 1, \ldots, d,$$

where we use hypothesis (b) to set $S^R_{\mathcal{M},j} := S^R_j|_{\mathcal{M}}$ for $j = 1, \ldots, d$, and choose the coefficient Hilbert space $\mathcal{U}$ so that

$$\dim \mathcal{U} = \operatorname{rank}(I - A_1^* A_1 - \cdots - A_d^* A_d).$$

By hypothesis (c) we may then choose the operator $C \colon \mathcal{M} \to \mathcal{U}$ so that

$$C^* C = I - A_1^* A_1 - \cdots A_d^* A_d.$$

Then $(C, \mathbf{A})$ is an isometric pair and, by hypothesis (d), $\mathbf{A}^*$ is strongly stable. Thus by part (2) of Proposition 2.3 it follows that the observability operator

$$\widehat{\mathcal{O}}_{C,\mathbf{A}} \colon f \mapsto C(I - Z(z)A)^{-1} f$$

is an isometry from $\mathcal{M}$ into $H^2_{\mathcal{U}}(\mathcal{F}_d)$. As observed for the general case in part (1) of Theorem 2.8, we have the intertwining condition

$$(S^R_j)^* \widehat{\mathcal{O}}_{C,\mathbf{A}} = \widehat{\mathcal{O}}_{C,\mathbf{A}} (S^R_{\mathcal{M},j})^*.$$

Taking adjoints then gives

$$(\widehat{\mathcal{O}}_{C,\mathbf{A}})^* S^R_j = S^R_{\mathcal{M},j} (\widehat{\mathcal{O}}_{C,\mathbf{A}})^*. \tag{2.62}$$

Let us set

$$\Theta = \iota \circ (\widehat{\mathcal{O}}_{C,\mathbf{A}})^{*} \colon H^2_{\mathcal{U}}(\mathcal{F}_d) \to H^2_{\mathcal{Y}}(\mathcal{F}_d),$$

where $\iota \colon \mathcal{M} \to H^2_{\mathcal{Y}}(\mathcal{F}_d)$ is the inclusion map. From hypothesis (a) that $\|\iota\| \le 1$, we see that $\|\Theta\| \le 1$. From the intertwining relation (2.62) (together with hypothesis (b)) it follows that

$$\Theta S_j^R = S_j^R \Theta$$

and it follows (see, e.g., Popescu (1995)) that Θ is a multiplication operator, i.e., there is a contractive multiplier $\theta \in \mathcal{S}_{nc,d}(\mathcal{U},\mathcal{Y})$ so that $\Theta = M_\theta$. From the fact that $\widehat{\mathcal{O}}_{C,\mathbf{A}} \colon \mathcal{M} \to H^2_{\mathcal{U}}(\mathcal{F}_d)$ is an isometry, it follows that $\mathrm{Ran}(\widehat{\mathcal{O}}_{C,\mathbf{A}})^* = \mathcal{M}$ and also that $\mathcal{M} = \theta \cdot H^2_{\mathcal{U}}(\mathcal{F}_d)$ with $\mathcal{M}$-norm given by (2.60). This completes the proof of necessity in statement (1) of Theorem 2.14 for the general case.

We now consider statement (2). In case $\mathcal{M}$ is isometrically included in $H^2_{\mathcal{Y}}(\mathcal{F}_d)$, for any $f \in H^2_{\mathcal{U}}(\mathcal{F}_d)$ we have

$$\|(\widehat{\mathcal{O}}_{C,\mathbf{A}})^* S_j^R f\|_{\mathcal{M}} = \|S_j^R (\widehat{\mathcal{O}}_{C,\mathbf{A}})^* f\|_{H^2_{\mathcal{Y}}(\mathcal{F}_d)} = \|(\widehat{\mathcal{O}}_{C,\mathbf{A}})^* f\|_{H^2_{\mathcal{U}}(\mathcal{F}_d)}$$

for $j = 1, \ldots, d$, since S_j^R is isometric on $H^2_{\mathcal{Y}}(\mathcal{F}_d)$. Since, as was observed above, $\widehat{\mathcal{O}}_{C,\mathbf{A}}$ is isometric, it follows that

$$\|P_{\mathrm{Ran}\, \widehat{\mathcal{O}}_{C,\mathbf{A}}} S_j f\| = \|f\| \quad \text{for all} \quad f \in \mathrm{Ran}\, \widehat{\mathcal{O}}_{C,\mathbf{A}}$$

and hence $\mathrm{Ran}\, \widehat{\mathcal{O}}_{C,\mathbf{A}}$ is invariant under S_j^R for $j = 1, \ldots, d$. As $\mathrm{Ran}\, \widehat{\mathcal{O}}_{C,\mathbf{A}}$ is also invariant under $(S_j^R)^*$ for each j by (2.62), we conclude that $\mathrm{Ran}\, \widehat{\mathcal{O}}_{C,\mathbf{A}}$ is reducing for $\mathbf{S}^R$. Since $\mathrm{Ran}\, C$ is dense in $\mathcal{U}$ by construction, we are now able to conclude that $\mathrm{Ran}\, \widehat{\mathcal{O}}_{C,\mathbf{A}}$ is all of $H^2_{\mathcal{U}}(\mathcal{F}_d)$ and hence $\widehat{\mathcal{O}}_{C,\mathbf{A}} \colon \mathcal{M} \to H^2_{\mathcal{U}}(\mathcal{F}_d)$ is actually unitary. It then follows finally that $\Theta = \iota \circ (\widehat{\mathcal{O}}_{C,\mathbf{A}})^*$ is isometric and hence θ is inner as asserted. This completes the proof of Theorem 2.14. □

A second application of these ideas is to operator model theory. For this application we are given only an operator-tuple $\mathbf{T} = (T_1, \ldots, T_d) \in \mathcal{L}(\mathcal{H})^d$ which is a row contraction, so $I - T_1 T_1^* - \cdots - T_d T_d^* \ge 0$. Set

$$D_{\mathbf{T}^*} := (I - T_1 T_1^* - \cdots - T_d T_d^*)^{1/2} \quad \text{and} \quad \mathcal{Y} := \mathrm{Ran}\, D_{\mathbf{T}^*}. \tag{2.63}$$

We apply the ideas of the previous sections concerning the general pair $(C, \mathbf{A})$ to a pair of the special form $(D_{\mathbf{T}^*}, \mathbf{T}^*)$. For simplicity we assume in addition that $\mathbf{T}^*$ is strongly stable, i.e.,

$$\lim_{N \to \infty} \sum_{v \in \mathcal{F}_d \colon |v| = N} \|\mathbf{T}^{*v} x\|^2 = 0 \quad \text{for all} \quad x \in \mathcal{H}.$$

Then we have the following dilation result.

Theorem 2.15 *Suppose that* $\mathbf{T} = (T_1, \ldots, T_d)$ *is a row contraction with* $\mathbf{T}^*$ *strongly stable as above and define the defect operator* $D_{\mathbf{T}^*}$ *and the coefficient space* $\mathcal{Y}$ *as in* (2.63). *Then there is a subspace* $\mathcal{M} \subset H^2_{\mathcal{Y}}(\mathcal{F}_d)$ *invariant for the backward shift operator-tuple* $\mathbf{S}^{R*}$ *on* $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ *so that* $\mathbf{T}$ *is unitarily equivalent to* $P_{\mathcal{M}} \mathbf{S}^R|_{\mathcal{M}}$. *In particular,* $\mathbf{T}$ *has a row-shift dilation unitarily equivalent to* $\mathbf{S}^R$ *on* $H^2_{\mathcal{Y}}(\mathcal{F}_d)$.

Proof By the same arguments as in the proof of Theorem 2.14, we see that

$$\widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}^*} \colon \mathcal{H} \to H^2_{\mathcal{Y}}(\mathcal{F}_d)$$

is isometric and satisfies the intertwining relations

$$(S_j^R)^* \widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}^*} = \widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}^*} T_j^* \quad \text{for } j = 1, \ldots, d.$$

If we then set

$$\mathcal{M} := \operatorname{Ran} \widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}^*},$$

then $\widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}*}$ implements the unitary equivalence between $\mathbf{T}$ and $P_{\mathcal{M}}\mathbf{S}|_{\mathcal{M}}$ as wanted. □

Remark 2.16 In the classical case $d = 1$, the procedure for constructing the unitary dilation of a contraction operator via the observability operator as in the proof of Theorem 2.15 corresponds to the construction of Douglas (1974) (see also Sz.-Nagy & Foiaş, 1970, Section I.10.1) which is an alternative to the more popular Schäffer-matrix construction of the unitary dilation (see Sz.-Nagy & Foiaş, 1970,Section I.5). Popescu (1989a, b) used an analogue of the Schäffer-matrix construction to construct the row-unitary dilation of a row-contraction operator-tuple. From the existence of this dilation, he went on to verify a von Neumann inequality (Popescu, 1991):

$$\|p(T_1, \ldots, T_d)\| \le \|p(S_1, \ldots, S_d)\|$$

for any polynomial $p \in \mathbb{C}\langle z\rangle$ in the noncommuting variable $z = (z_1, \ldots, z_d)$. He returned to this topic in Popescu (1999) to give another proof of the von Neumann inequality (actually a more general version involving nonanalytic polynomials) based on the Poisson transform: for $\mathbf{T}$ a *strict* row-contraction (one can reduce the general case of a row-contraction to the case of a strict row-contraction via a limiting procedure), one defines the Poisson transform $P(\mathbf{T})$: $\mathcal{L}(H^2_{\mathcal{Y}}(\mathcal{F}_d), \mathcal{H}) \to \mathcal{L}(\mathcal{H})$ by

$$P(\mathbf{T})[X] = (\widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}^*})^* X \widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}^*}. \tag{2.64}$$

It is argued in Popescu (1999) (as well as in Chalendar (2003) in the context of the classical case) that this is an elementary (i.e., dilation-free) proof of the von Neumann inequality. Indeed, as argued in Chalendar (2003), this proof of the von Neumann inequality goes back to the paper of Heinz (1952). However, we would argue that the dilation is very near the surface in this proof as well, since the Poisson kernel, i.e., the observability operator $\widehat{\mathcal{O}}_{D_{\mathbf{T}^*},\mathbf{T}^*}$, provides the factorization of the Poisson transform (2.64) and is also the operator embedding the state space $\mathcal{H}$ into the dilation space $H^2_{\mathcal{Y}}(\mathcal{F}_d)$ in the Douglas approach to dilation theory.

3 The commutative-variable Arveson-space setting

3.1 Output stability and Stein equations: the commutative-variable case

To introduce the commutative multidimensional counterpart of the Hardy space $H^2(\mathbb{D})$, we recall standard multivariable notations: for a multi-integer

$$\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d_+$$

and a point $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d) \in \mathbb{C}^d$, we set $|\mathbf{n}| = n_1 + n_2 + \cdots + n_d$, $\mathbf{n}! = n_1! n_2! \ldots n_d!$ and

$$\boldsymbol{\lambda}^{\mathbf{n}} = \lambda_1^{n_1} \lambda_2^{n_2} \ldots \lambda_d^{n_d}. \tag{3.1}$$

The space can be derived from the full Fock space by "letting the variables commute." For this purpose we introduce the abelianization map $\mathbf{a}$: $\mathcal{F}_d \to \mathbb{Z}_+^d$ given by

$$\mathbf{a}(i_N \cdots i_1) = (n_1, \ldots, n_d), \quad \text{where} \quad n_k = \#\{\ell\colon\ i_\ell = k\} \ \text{ for } \ k = 1, \ldots, d.$$

A key combinatorial fact is that

$$\#\mathbf{a}^{-1}(\mathbf{n}) = \frac{|\mathbf{n}|!}{\mathbf{n}!}. \tag{3.2}$$

We then consider the symmetric Fock space $\ell^2_{\mathcal{Y}}(\mathcal{SF}_d)$ equal to the subspace of $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$ spanned by the elements $\chi_{\mathbf{n}} y$ ($n \in \mathbb{Z}_+^d$ and $y \in \mathcal{Y}$) where $\chi_{\mathbf{n}}$ is given by

$$\chi_{\mathbf{n}} = \sum_{v\colon \mathbf{a}(v) = \mathbf{n}} \chi_v,$$

where χ_v is defined as just below (2.2). Note that

$$\|\chi_{\mathbf{n}}\|^2_{\ell^2_{\mathbb{C}}(\mathcal{F}_d)} = \sum_{v \in \mathcal{F}_d\colon \mathbf{a}(v) = n} 1 = \frac{|\mathbf{n}|!}{\mathbf{n}!}$$

and hence, if $\mathcal{B}$ is an orthonormal basis for $\mathcal{Y}$, then an orthonormal basis for $\ell^2_{\mathcal{Y}}(\mathcal{SF}_d)$ is the set

$$\left\{ \sqrt{\frac{\mathbf{n}!}{|\mathbf{n}|!}} \chi_{\mathbf{n}} y\colon \mathbf{n} \in \mathbb{Z}_+^d,\ y \in \mathcal{B} \right\}.$$

It is then natural to identify $\ell^2_{\mathcal{Y}}(\mathcal{SF}_d)$ with the weighted sequence space $\ell^2_{w,\mathcal{Y}}(\mathbb{Z}_+^d)$ consisting of all $\mathcal{Y}$-valued $\mathbb{Z}_+^d$-indexed sequences $\{f_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{Z}_+^d}$ for which the norm given by

$$\|\{f_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{Z}_+^d}\|_{\ell^2_{w,\mathcal{Y}}(\mathbb{Z}_+^d)} = \sum_{\mathbf{n} \in \mathbb{Z}_+^d} w(\mathbf{n}) \|f_{\mathbf{n}}\|^2, \quad \text{where } \ w(\mathbf{n}) = \frac{\mathbf{n}!}{|\mathbf{n}|!},$$

is finite. We abbreviate $\ell^2_{w,\mathbb{C}}(\mathbb{Z}_+^d)$ to $\ell^2_w(\mathbb{Z}_+^d)$ and observe that

$$\ell^2_{w,\mathcal{Y}}(\mathbb{Z}_+^d) = \ell^2_w(\mathbb{Z}_+^d) \otimes \mathcal{Y}. \tag{3.3}$$

The commutative d-variable Z-transform

$$\{f_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{Z}_+^d} \mapsto \widehat{f}^{\mathbf{a}}(\boldsymbol{\lambda}) = \sum_{\mathbf{n} \in \mathbb{Z}_+^d} f_{\mathbf{n}} \boldsymbol{\lambda}^{\mathbf{n}}$$

maps $\ell^2_w(\mathbb{Z}_+^d)$ unitarily onto the *Arveson space*

$$\mathcal{H}(k_d) := \left\{ f(\boldsymbol{\lambda}) = \sum_{\mathbf{n} \in \mathbb{Z}_+^d} f_{\mathbf{n}} \boldsymbol{\lambda}^{\mathbf{n}} :\ \|f\|^2 = \sum_{\mathbf{n} \in \mathbb{Z}_+^d} \frac{\mathbf{n}!}{|\mathbf{n}|!} \cdot |f_{\mathbf{n}}|^2 < \infty \right\}$$

with inner product given by

$$\langle f, g\rangle_{\mathcal{H}(k_d)} = \sum_{\mathbf{n}\in\mathbb{Z}_+^d} \frac{\mathbf{n}!}{|\mathbf{n}|!} f_n \overline{g_n}$$

if

$$f(\boldsymbol{\lambda}) = \sum_{\mathbf{n}\in\mathbb{Z}_+^d} f_{\mathbf{n}} \boldsymbol{\lambda}^{\mathbf{n}} \quad \text{and} \quad g(\boldsymbol{\lambda}) = \sum_{\mathbf{n}\in\mathbb{Z}_+^d} g_{\mathbf{n}} \boldsymbol{\lambda}^{\mathbf{n}}.$$

Then it follows that the set $\{\sqrt{\frac{|\mathbf{n}|!}{\mathbf{n}!}}\boldsymbol{\lambda}^{\mathbf{n}} \colon \mathbf{n} \in \mathbb{Z}_+^d\}$ is an orthonormal basis for $\ell^2_w(\mathbb{Z}_+^d)$. By general principles concerning reproducing kernel Hilbert spaces we see that $\mathcal{H}(k_d)$ is a reproducing kernel Hilbert space of functions analytic on the unit ball

$$\mathbb{B}^d = \left\{ \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d) \in \mathbb{C}^d \colon \sum_{k=1}^d |\lambda_k|^2 < 1 \right\}$$

with reproducing kernel $k_d(\boldsymbol{\lambda}, \boldsymbol{\zeta})$ given by

$$k_d(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = \sum_{\mathbf{n}\in\mathbb{Z}_+^d} \frac{|\mathbf{n}|!}{\mathbf{n}!} \boldsymbol{\lambda}^{\mathbf{n}} \overline{\boldsymbol{\zeta}}^{\mathbf{n}} = \sum_{n=0}^{\infty} \left(\lambda_1\overline{\zeta_1} + \cdots + \lambda_d\overline{\zeta_d}\right)^n = \frac{1}{1 - \langle \boldsymbol{\lambda}, \boldsymbol{\zeta}\rangle}$$

(see, e.g., Arveson, (1998)). This justifies the notation $\mathcal{H}(k_d)$ for the space. In analogy to (3.3), we will use notation $\mathcal{H}_{\mathcal{Y}}(k_d) := \mathcal{H}(k_d) \otimes \mathcal{Y}$ for the tensor product Hilbert space that is characterized by

$$\mathcal{H}_{\mathcal{Y}}(k_d) = \left\{ f(\boldsymbol{\lambda}) = \sum_{\mathbf{n}\in\mathbb{Z}_+^d} f_{\mathbf{n}} \boldsymbol{\lambda}^{\mathbf{n}} : \|f\|^2 = \sum_{\mathbf{n}\in\mathbb{Z}_+^d} \frac{\mathbf{n}!}{|\mathbf{n}|!} \cdot \|f_{\mathbf{n}}\|^2_{\mathcal{Y}} < \infty \right\}.$$

If we define the map Π by

$$\Pi \colon \{f_v\}_{v\in\mathcal{F}_d} \mapsto \left\{ \sum_{v\colon \mathbf{a}(v)=\mathbf{n}} f_v \right\}_{\mathbf{n}\in\mathbb{Z}_+^d} \tag{3.4}$$

then each basis vector $\chi_v \in \ell^2(\mathcal{F}_d)$ ($v \in \mathcal{F}_d$) is mapped via Π to its abelianization $\chi_{\mathbf{n}} \in \ell^2_w(\mathbb{Z}_+^d)$ and then Π is extended to the whole space $\ell^2(\mathcal{F}_d)$ via linearity. The norm on $\ell^2_w(\mathbb{Z}_+^d)$ is arranged so as to make Π a coisometry from $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$ onto $\ell^2_{w,\mathcal{Y}}(\mathbb{Z}_+^d)$ with initial space equal to $\ell^2_{\mathcal{Y}}(\mathcal{SF}_d)$ and with kernel equal to the subspace $\ell^2_{\mathcal{Y}}(\mathcal{SF}_d)^\perp$ of $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$ given by

$$\ell^2_{\mathcal{Y}}(\mathcal{SF}_d)^\perp = \left\{ \{f_v\}_{v\in\mathcal{F}_d} \colon \sum_{v\in\mathcal{F}_d\colon \mathbf{a}(v)=\mathbf{n}} f_v = 0 \ \text{ for each } \ \mathbf{n} \in \mathbb{Z}_+^d \right\}.$$

If we introduce the Z-transformed version $\widehat{\Pi}\colon H^2(\mathcal{F}_d) \to \mathcal{H}(k_d)$ via

$$\widehat{\Pi}\colon \sum_{v\in\mathcal{F}_d} f_v z^v \mapsto \sum_{\mathbf{n}\in\mathbb{Z}_+^d} \left[\sum_{v\in\mathcal{F}_d\colon \mathbf{a}(v)=\mathbf{n}} f_v \right] \boldsymbol{\lambda}^{\mathbf{n}}$$

then similarly $\widehat{\Pi}$ is a coisometry from $\mathcal{H}^2(\mathcal{F}_d)$ onto $\mathcal{H}(k_d)$ with initial space equal to the subspace

$$H^2(\mathcal{SF}_d) := \left\{ \sum_{v \in \mathcal{F}_d} f_{\mathbf{a}(v)} z^v \colon \sum_{\mathbf{n} \in \mathbb{Z}_+^d} |f_{\mathbf{n}}|^2 < \infty \right\}$$

with kernel equal to

$$H^2(\mathcal{SF}_d)^\perp = \left\{ \sum_{v \in \mathcal{F}_d} f_v z^v \in H^2(\mathcal{F}_d) \colon \sum_{v \colon \mathbf{a}(v) = \mathbf{n}} f_v = 0 \text{ for each } \mathbf{n} \in \mathbb{Z}_+^d \right\}.$$

This gives the natural link between the Fock-space norm on formal power series and the Arveson-space norm on analytic functions on the unit ball and is the basis for the application of noncommutative results to prove commutative results in Arias and Popescu (2000), Davidson and Pitts (1998), and Popescu (1998).

By a commutative d-dimensional linear system we mean a linear system with evolution along the integer lattice $\mathbb{Z}_+^d$ rather than along the free semigroup $\mathcal{F}_d$. A particular type of such a system is a system of the Fornasini–Marchesini form given by (1.18). If we specify an initial condition $x(0) = x^0 \in \mathcal{X}$ along with an input sequence $\{u^0(\mathbf{n})\}_{\mathbf{n} \in \mathbb{Z}_+^d}$ and impose the boundary conditions that $x(\mathbf{n}) = 0$ whenever $\mathbf{n}$ is outside the positive orthant $\mathbb{Z}_+^d$, then the system equations uniquely determine a full system trajectory $\{u(\mathbf{n}), x(\mathbf{n}), y(\mathbf{n})\}$ consistent with $x(0) = x^0$ and $u(\mathbf{n}) = u^0(\mathbf{n})$ for $\mathbf{n} \in \mathbb{Z}_+^d$.

If Π is the projection map introduced in (3.4) formally extended to be defined on all $\mathcal{F}_d$-indexed sequences to generate a $\mathbb{Z}_+^d$-indexed sequence

$$\Pi\colon \{u(v)\}_{v \in \mathcal{F}_d} \mapsto \left\{ \sum_{v \colon \mathbf{a}(v) = \mathbf{n}} u(v) \right\}_{\mathbf{n} \in \mathbb{Z}_+^d},$$

$$\Pi\colon \{x(v)\}_{v \in \mathcal{F}_d} \mapsto \left\{ \sum_{v \colon \mathbf{a}(v) = \mathbf{n}} x(v) \right\}_{\mathbf{n} \in \mathbb{Z}_+^d},$$

$$\Pi\colon \{y(v)\}_{v \in \mathcal{F}_d} \mapsto \left\{ \sum_{v \colon \mathbf{a}(v) = \mathbf{n}} y(v) \right\}_{\mathbf{n} \in \mathbb{Z}_+^d}$$

then one can check the claim: $\{(\Pi u)(\mathbf{n}), (\Pi x)(\mathbf{n}), (\Pi y)(\mathbf{n})\}_{\mathbf{n} \in \mathbb{Z}_+^d}$ *satisfies the system equations* (1.18) *whenever* $\{u(v), x(v), y(v)\}_{v \in \mathcal{F}_d}$ *satisfies the system equations* (1.13). Indeed the first system equation in (1.13) can be rewritten in the form

$$x(v) = \sum_{k=1}^{d} A_k x(k^{-1}v) + \sum_{k=1}^{d} B_k u(k^{-1}v).$$

Here we use the convention that

$$k^{-1}v = \begin{cases} v', & \text{if } v = kv', \\ \text{undefined}, & \text{otherwise} \end{cases}$$

for v a word in $\mathcal{F}_d$ and $k \in \{1, \ldots, d\}$ a letter and that $x(k^{-1}v)$ is interpreted to be 0 if $k^{-1}v$ is undefined. Summing over $v \in \mathcal{F}_d$ with $\mathbf{a}(v) = \mathbf{n}$ then gives

$$(\Pi x)(\mathbf{n}) = \sum_{k=1}^{d} A_k \sum_{v:\, \mathbf{a}(v)=\mathbf{n}} x(k^{-1}v) + \sum_{k=1}^{d} B_k \sum_{v:\, \mathbf{a}(v)=\mathbf{n}} u(k^{-1}v).$$

Now observe that

$$\{k^{-1}v \colon \mathbf{a}(v) = \mathbf{n}\} = \{v' \in \mathcal{F}_d \colon \mathbf{a}(v') = \mathbf{n} - e_k\}$$

and arrive at

$$(\Pi x)(\mathbf{n}) = \sum_{k=1}^{d} A_k (\Pi x)(\mathbf{n} - e_k) + \sum_{k=1}^{d} B_k (\Pi u)(\mathbf{n} - e_k).$$

We see that $\{(\Pi u)(\mathbf{n}), (\Pi x)(\mathbf{n}), (\Pi y)(\mathbf{n})\}$ satisfies the first of the system equations (1.18). That $\{(\Pi u)(\mathbf{n}), (\Pi x)(\mathbf{n}), (\Pi y)(\mathbf{n})\}$ satisfies the second system equation in (1.18) is a simple consequence of linearity. Conversely, given a trajectory $\{u(\mathbf{n}), x(\mathbf{n}), y(\mathbf{n})\}_{\mathbf{n} \in \mathbb{Z}^d_+}$ of (1.18), let $\{u_\ell(v)\}_{v \in \mathcal{F}_d}$ be any $\mathcal{U}$-valued $\mathcal{F}_d$-indexed sequence such that $\Pi u_\ell = u$ and set $x_\ell(\emptyset) = x(0)$. Then the noncommutative system equations (1.13) recursively uniquely determine a full system trajectory $\{u_\ell(v), x_\ell(v), y_\ell(v)\}_{v \in \mathcal{F}_d}$ of (1.13) with this pre-assigned input string and initial condition. By the claim verified above, it follows that $(\Pi u_\ell, \Pi x_\ell, \Pi y_\ell)$ is again a system trajectory. By the uniqueness of solution of the initial-value problem for the system (1.18), it follows that $\{(\Pi u, \Pi x, \Pi y)\} = \{(u, x, y)\}$. Thus, any trajectory $\{u(\mathbf{n}), x(\mathbf{n}), y(\mathbf{n})\}_{\mathbf{n} \in \mathbb{Z}^d_+}$ can be lifted to a trajectory $\{u_\ell(v), x_\ell(v), y_\ell(v)\}_{v \in \mathcal{F}_d}$ of (1.13), i.e., $\{u_\ell(v), x_\ell(v), y_\ell(v)\}_{v \in \mathcal{F}_d}$ is a trajectory of (1.13) such that

$$\{\Pi u_\ell, \Pi x_\ell, \Pi y_\ell\} = \{u, x, y\}.$$

In this way, we view the Fornasini–Marchesini commutative system (1.18) as the abelianization of the noncommutative Fornasini–Marchesini system (1.13).

Since the commutative Fornasini–Marchesini system (1.18) is just the abelianization of the noncommutative Fornasini–Marchesini system (1.13) and we have already derived the formula (2.6) for the solution of the noncommutative initial-value problem, we see that the solution of the initial-value problem for the commutative Fornasini–Marchesini system (1.18) is simply the abelianization of the corresponding formula for the noncommutative case:

$$(\widehat{\Pi y})(\boldsymbol{\lambda}) = C(I - Z(\boldsymbol{\lambda})A)^{-1}x(0) + T_\Sigma(\boldsymbol{\lambda}) \cdot (\widehat{\Pi u})(\boldsymbol{\lambda}), \tag{3.5}$$

where the *transfer function* $T_\Sigma(\boldsymbol{\lambda})$ for the commutative Fornasini–Marchesini system is given by

$$T_\Sigma(\boldsymbol{\lambda}) = D + C(I - Z(\boldsymbol{\lambda})A)^{-1}Z(\boldsymbol{\lambda})B.$$

This gives a derivation of the transfer function relationship (3.5) (via the connection with noncommutative systems) which is an alternative to the usual direct approach via commutative multivariable Z-transform (see, e.g., Ball, Sadosky, & Vinnikov, 2005).

The zero input string simplifies the system to

$$\begin{aligned} x(\mathbf{n}) &= A_1 x(\mathbf{n} - e_1) + \cdots + A_d x(\mathbf{n} - e_d), \\ y(\mathbf{n}) &= C x(\mathbf{n}). \end{aligned} \tag{3.6}$$

Given a pair $(C, \mathbf{A})$, we have the option of considering $(C, \mathbf{A})$ as coming from a noncommutative or a commutative system. If we consider the associated noncommutative system, the output string associated with initial state $x(\emptyset) = x$ (and zero input string) is the $\mathcal{Y}$-valued function on $\mathcal{F}_d$ given by

$$\mathcal{O}_{C,\mathbf{A}} x = \{C\mathbf{A}^v x\}_{v \in \mathcal{F}_d}$$

and $(C, \mathbf{A})$ is considered output stable if this output string is in $\ell^2_{\mathcal{Y}}(\mathcal{F}_d)$ for all $x \in \mathcal{H}$. We say that the commutative system (3.6) is *output stable* (and in this case we will say that the pair $(C, \mathbf{A})$ is $\mathbf{a}$*-output stable*) if $\Pi(\mathcal{O}_{C,\mathbf{A}} x) \in \ell^2_{w,\mathcal{Y}}(\mathbb{Z}^d_+)$ for all $x \in \mathcal{H}$, or equivalently, if $\widehat{\Pi}\widehat{\mathcal{O}}_{C,\mathbf{A}} x$ is in the Arveson space $\mathcal{H}_{\mathcal{Y}}(k_d)$ for all choices of initial state $x \in \mathcal{X}$. We note that $\widehat{\Pi}\widehat{\mathcal{O}}_{C,\mathbf{A}} x$ can be computed explicitly as

$$(\widehat{\Pi}\widehat{\mathcal{O}}_{C,\mathbf{A}} x)(\boldsymbol{\lambda}) = C(I - Z(\boldsymbol{\lambda})A)^{-1} x.$$

Thus another equivalent formulation of $\mathbf{a}$-output stability is:

Definition 3.1 A pair $(C, \mathbf{A})$ is $\mathbf{a}$-output stable means that the function $C(I - Z(\boldsymbol{\lambda})A)^{-1} x$ belongs to $\mathcal{H}_{\mathcal{Y}}(k_d)$ for every $x \in \mathcal{H}$, or equivalently (by the closed graph theorem), the operator $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ from $\mathcal{X}$ to $\mathcal{H}_{\mathcal{Y}}(k_d)$ defined by

$$\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} = \widehat{\Pi}\widehat{\mathcal{O}}_{C,\mathbf{A}} \colon x \mapsto C(I - Z(\boldsymbol{\lambda})A)^{-1} x = \sum_{\mathbf{n} \in \mathbb{Z}^d_+} \left(\sum_{v \in \mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^v x \right) \boldsymbol{\lambda}^{\mathbf{n}} \tag{3.7}$$

is bounded.

The inverse Z-transform sends the function

$$\widehat{y}^{\mathbf{a}}(\boldsymbol{\lambda}) = C(I - Z(\boldsymbol{\lambda})A)^{-1} x = \sum_{\mathbf{n} \in \mathbb{Z}^d_+} \left(\sum_{v \in \mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^v x \right) \boldsymbol{\lambda}^{\mathbf{n}}$$

to the string $\{y(\mathbf{n})\}_{\mathbf{n} \in \mathbb{Z}^d_+}$ with

$$y(\mathbf{n}) = \sum_{v \in \mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^v x, \quad \mathbf{n} \in \mathbb{Z}^d_+ \tag{3.8}$$

and $\widehat{y}^{\mathbf{a}}$ belongs to $\mathcal{H}_{\mathcal{Y}}(k_d)$ if and only if $\{y(\mathbf{n})\}_{\mathbf{n} \in \mathbb{Z}^d_+} \in \ell^2_{w,\mathcal{Y}}(\mathbb{Z}^d_+)$. Thus, the operator $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ introduced in (3.7) is the Z-transformed version of the observability operator

$$\mathcal{O}^{\mathbf{a}}_{C,\mathbf{A}} \colon \; x \mapsto \left\{ \sum_{v \in \mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^v x \right\}_{\mathbf{n} \in \mathbb{Z}^d_+} \tag{3.9}$$

and a pair $(C, \mathbf{A})$ is $\mathbf{a}$-output stable if and only if $\mathcal{O}^{\mathbf{a}}_{C,\mathbf{A}}$ is bounded as an operator from $\mathcal{X}$ into $\ell^2_{w,\mathcal{Y}}(\mathbb{Z}^d_+)$. In this case it makes sense to introduce the *observability gramian*

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} := (\mathcal{O}^{\mathbf{a}}_{C,\mathbf{A}})^* \mathcal{O}^{\mathbf{a}}_{C,\mathbf{A}} = (\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}})^* \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$$

and its representation in terms of strongly converging series

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \sum_{\mathbf{n} \in \mathbb{Z}^d_+} \frac{\mathbf{n}!}{|\mathbf{n}|!} \left(\sum_{v,u \in \mathbf{a}^{-1}(\mathbf{n})} \mathbf{A}^{*v^\top} C^* C\mathbf{A}^u \right) \tag{3.10}$$

follows immediately by definitions (3.9), (3.7) and the formulas for the inner products in $\ell^2_{w,\mathcal{Y}}(\mathbb{Z}^d_+)$ and $\mathcal{H}_{\mathcal{Y}}(k_d)$.

Definition 3.2 We say that the pair $(C, \mathbf{A})$ is $\mathbf{a}$-*observable* if $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ is positive-definite and *exactly* $\mathbf{a}$-*observable* if $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ is strictly positive definite.

By Theorem 2.2 (2) we know that the observability gramian $\mathcal{G}_{C,\mathbf{A}}$ satisfies the Stein equation (2.15). It turns out that the abelianized observability gramian $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ satisfies a reverse Stein inequality (the reverse of (2.14)).

Proposition 3.3 *Let $(C, \mathbf{A})$ be an $\mathbf{a}$-output-stable pair and let $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ be the abelianized observability gramian* (3.10)*. Then $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ satisfies the reverse Stein inequality*

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} - A_1^*\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}A_1 - \cdots - A_d^*\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}A_d \leq C^*C. \tag{3.11}$$

Moreover, the following are equivalent:

(1) *Equality holds in* (3.11).
(2) $\mathbf{A}$ *is C-abelian in the sense that*

$$C\mathbf{A}^v = C\mathbf{A}^u, \quad \textit{whenever } v, u \in \mathcal{F}_d \textit{ and } \mathbf{a}(v) = \mathbf{a}(u). \tag{3.12}$$

(3) *The observability gramian and the abelianized observability gramian are identical:*

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}}.$$

Proof It suffices to show that the operator Q given by

$$Q := C^*C - \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} + \sum_{j=1}^{d} A_j^*\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}A_j \tag{3.13}$$

is positive semidefinite. To this end, plug (3.10) into (3.13) to get

$$Q = \sum_{N=1}^{\infty} Q_N, \tag{3.14}$$

where Q_N is given by

$$Q_N = \sum_{j=1}^{d} A_j^* \left[\sum_{\mathbf{m}\in\mathbb{Z}^d_+ :\ |\mathbf{m}|=N-1} \frac{\mathbf{m}!}{(N-1)!} \sum_{v,u\in\mathbf{a}^{-1}(\mathbf{m})} \mathbf{A}^{*v^\top} C^*C\mathbf{A}^u \right] A_j \\ - \sum_{\mathbf{n}\in\mathbb{Z}^d_+ :\ |\mathbf{n}|=N} \frac{\mathbf{n}!}{N!} \sum_{v,u\in\mathbf{a}^{-1}(\mathbf{n})} \mathbf{A}^{*v^\top} C^*C\mathbf{A}^u. \tag{3.15}$$

We introduce the notation

$$W(\mathbf{n}) = \sum_{u\in\mathbf{a}^{-1}(\mathbf{n})} \mathbf{A}^u \quad (\mathbf{n} \in \mathbb{Z}^d_+) \tag{3.16}$$

and extend the notation to the all of $\mathbb{Z}^d$ by

$$W(\mathbf{n}) = 0 \text{ if } \mathbf{n} \in \mathbb{Z}^d \setminus \mathbb{Z}^d_+. \tag{3.17}$$

With these definitions we have the equality

$$W(\mathbf{n}) = \sum_{i=1}^{d} W(\mathbf{n} - e_i)A_i \qquad \text{for every } \mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}_+^d, \tag{3.18}$$

where $e_1, \ldots, e_d \in \mathbb{Z}_+^d$ are defined in (1.19). Write formula (3.15) in terms of (3.16) as

$$\begin{aligned} Q_N = & \sum_{j=1}^{d} A_j^* \left[\sum_{\mathbf{m} \in \mathbb{Z}_+^d \colon |\mathbf{m}| = N-1} \frac{\mathbf{m}!}{(N-1)!} W(\mathbf{m})^* C^* C W(\mathbf{m}) \right] A_j \\ & - \sum_{\mathbf{n} \in \mathbb{Z}_+^d \colon |\mathbf{n}| = N} \frac{\mathbf{n}!}{N!} \cdot W(\mathbf{n})^* C^* C W(\mathbf{n}). \end{aligned} \tag{3.19}$$

Upon rearranging the terms in the first series in (3.19) and substituting formula (3.18) into the second series, we arrive at

$$\begin{aligned} Q_N = & \sum_{\mathbf{n} \in \mathbb{Z}_+^d \colon |\mathbf{n}| = N} \sum_{j=1}^{d} \frac{(\mathbf{n} - e_j)!}{(N-1)!} A_j^* W(\mathbf{n} - e_j)^* C^* C W(\mathbf{n} - e_j) A_j \\ & - \sum_{\mathbf{n} \in \mathbb{Z}_+^d \colon |\mathbf{n}| = N} \left(\sum_{i,j=1}^{d} \frac{\mathbf{n}!}{N!} A_i^* W(\mathbf{n} - e_i)^* C^* C W(\mathbf{n} - e_j) A_j \right). \end{aligned} \tag{3.20}$$

We now consider the terms in (3.20) that correspond to a fixed $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}_+^d$ (with $|\mathbf{n}| = N$). Denoting the sum of these terms by $S_{\mathbf{n}}$ we have

$$\begin{aligned} S_{\mathbf{n}} = & \sum_{j=1}^{d} \frac{(\mathbf{n} - e_j)!}{(N-1)!} A_j^* W(\mathbf{n} - e_j)^* C^* C W(\mathbf{n} - e_j) A_j \\ & - \frac{\mathbf{n}!}{N!} \sum_{i,j=1}^{d} A_i^* W(\mathbf{n} - e_i)^* C^* C W(\mathbf{n} - e_j) A_j \\ = & \sum_{j=1}^{d} \frac{(\mathbf{n} - e_j)!}{N!} (N - n_j) A_j^* W(\mathbf{n} - e_j)^* C^* C W(\mathbf{n} - e_j) A_j \\ & - \frac{\mathbf{n}!}{N!} \sum_{i,j \in \{1,\ldots,d\} \colon i \neq j} A_i^* W(\mathbf{n} - e_i)^* C^* C W(\mathbf{n} - e_j) A_j. \end{aligned} \tag{3.21}$$

Note that by convention (3.17), the indices i and j in the latter summations vary on the set

$$\mathcal{I}_{\mathbf{n}} = \{\ell \in \{1, \ldots, d\} : \ n_\ell > 0\}$$

rather than $\{1, \ldots, d\}$. Furthermore, since

$$N - n_j = |\mathbf{n}| - n_j = \sum_{i \in \mathcal{I}_{\mathbf{n}} \colon i \neq j} n_i$$

and

$$(\mathbf{n} - e_j)! = (\mathbf{n} - e_j - e_i)!\, n_i \quad (i \neq j)$$

one can rewrite the first sum on the right hand side in (3.21) as

$$\sum_{i,j\in\mathcal{I}_{\mathbf{n}}:\, i\neq j} \frac{(\mathbf{n}-e_j)!}{N!}\, n_i A_j^* W(\mathbf{n}-e_j)^* C^* C W(\mathbf{n}-e_j) A_j$$

$$= \sum_{i,j\in\mathcal{I}_{\mathbf{n}}:\, i\neq j} \frac{(\mathbf{n}-e_j-e_i)!}{N!}\, n_i^2 A_j^* W(\mathbf{n}-e_j)^* C^* C W(\mathbf{n}-e_j) A_j.$$

Plugging this into the right hand side in (3.21) leads us to

$$S_{\mathbf{n}} = \sum_{i,j\in\mathcal{I}_{\mathbf{n}}: i\neq j} \frac{(\mathbf{n}-e_j-e_i)!}{N!}\Big[n_i^2 A_j^* W(\mathbf{n}-e_j)^* C^* C W(\mathbf{n}-e_j) A_j$$

$$-n_i n_j A_i^* W(\mathbf{n}-e_i)^* C^* C W(\mathbf{n}-e_j) A_j\Big]$$

$$= \sum_{i,j\in\mathcal{I}_{\mathbf{n}}: i\neq j} \frac{1}{2}\frac{(\mathbf{n}-e_j-e_i)!}{N!} R_{\mathbf{n},i,j}^* R_{\mathbf{n},i,j}, \tag{3.22}$$

where

$$R_{\mathbf{n},i,j} = C\left[n_i W(\mathbf{n}-e_j)A_j - n_j W(\mathbf{n}-e_i)A_i\right]. \tag{3.23}$$

Representation (3.22) implies that $S_{\mathbf{n}}$ is positive semidefinite and therefore $Q_N \geq 0$ for every $N \in \mathbb{N}$. By (3.14), the operator Q defined in (3.13) is positive semidefinite which completes the proof of (3.11).

We now show the equivalence of (1), (2) and (3) in the second part of Proposition 3.3.

Proof of (1) $\Longrightarrow$ (2) Assume condition (1), i.e., that the reverse Stein inequality (3.11) is satisfied with equality. Then representation (3.22) implies that $R_{\mathbf{n},i,j} = 0$ for all $\mathbf{n} \in \mathbb{Z}_+^d$. By (3.23), this means that

$$n_i C W(\mathbf{n}-e_j)A_j = n_j C W(\mathbf{n}-e_i)A_i \qquad (\mathbf{n} \in \mathbb{Z}_+^d). \tag{3.24}$$

Now we shall prove (3.12) by induction (on the length of words $v, u \in \mathcal{F}_d$). The basis of induction ($|v| = |u| = 0$) is trivial. Assume that (3.12) holds true, whenever $|v| = |u| < N$. Then in particular, we have for every $\mathbf{m} \in \mathbb{Z}_+^d$ with $|\mathbf{m}| < N$:

$$CW(\mathbf{m}) = \sum_{w\in\mathbf{a}^{-1}(\mathbf{m})} C\mathbf{A}^w = \frac{|\mathbf{m}|!}{\mathbf{m}!} C\mathbf{A}^{w_0} \quad \text{for every } \ w_0 \in \mathbf{a}^{-1}(\mathbf{m}). \tag{3.25}$$

Now take two words $v, u \in \mathcal{F}_d$ of the length N and let

$$\mathbf{a}(v) = \mathbf{a}(u) =: \mathbf{n} = (n_1, \ldots, n_d). \tag{3.26}$$

If $v = \widetilde{v}i$ and $u = \widetilde{u}i$ for some $\widetilde{v}, \widetilde{u} \in \mathcal{F}_d$ and $i \in \{1, \ldots, d\}$, then we have $C\mathbf{A}^{\widetilde{v}} = C\mathbf{A}^{\widetilde{u}}$ by the induction hypothesis and therefore,

$$C\mathbf{A}^v = C\mathbf{A}^{\widetilde{v}}A_i = C\mathbf{A}^{\widetilde{u}}A_i = C\mathbf{A}^u.$$

Let $v = \widetilde{v}i$ and $u = \widetilde{u}j$ for some $i, j \in \{1, \ldots, d\}$ and $i \neq j$. By (3.26), $\mathbf{a}(\widetilde{v}) = \mathbf{n} - e_i$ and $\mathbf{a}(\widetilde{u}) = \mathbf{n} - e_j$. By (3.25), we have

$$CW(\mathbf{n}-e_j) = \frac{(N-1)!}{(\mathbf{n}-e_j)!} C\mathbf{A}^{\widetilde{v}}, \tag{3.27}$$

$$CW(\mathbf{n}-e_i) = \frac{(N-1)!}{(\mathbf{n}-e_i)!} C\mathbf{A}^{\widetilde{u}}. \tag{3.28}$$

Multiplying (3.27) and (3.28) on the right by n_iA_j and n_jA_i, respectively, we get

$$n_iCW(\mathbf{n}-e_j)A_j = n_i\frac{(N-1)!}{(\mathbf{n}-e_j)!}C\mathbf{A}^vA_j = n_i\frac{(N-1)!}{(\mathbf{n}-e_j)!}C\mathbf{A}^{vj} = n_in_j\frac{(N-1)!}{\mathbf{n}!}C\mathbf{A}^v$$

and

$$n_jCW(\mathbf{n}-e_i)A_i = n_j\frac{(N-1)!}{(\mathbf{n}-e_i)!}C\mathbf{A}^vA_i = n_j\frac{(N-1)!}{(\mathbf{n}-e_i)!}C\mathbf{A}^{ui} = n_jn_i\frac{(N-1)!}{\mathbf{n}!}Y\mathbf{A}^u.$$

By (3.24), the left hand side expressions in the two latter equalities are equal. Upon comparing the right hand side expressions we get $C\mathbf{A}^v = C\mathbf{A}^u$, i.e., $\mathbf{A}$ is C-abelian as wanted.

Proof of (2) $\Longrightarrow$ (3) Assume now that $\mathbf{A}$ is C-abelian, i.e., that (3.12) holds. Then the identify $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}}$ is an immediate consequence of the series representations (3.10) and (2.10) for $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ and $\mathcal{G}_{C,\mathbf{A}}$ respectively.

Proof of (3) $\Longrightarrow$ (1) We know from Theorem 2.8 (2) that $\mathcal{G}_{C,\mathbf{A}}$ satisfies the Stein equation, i.e., $\mathcal{G}_{C,\mathbf{A}}$ satisfies the Stein inequality (3.11) with equality. Hence trivially $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ satisfies (3.11) with equality whenever $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}}$. This completes the proof of Proposition 3.3. □

Example 3.4 If $(C,\mathbf{A})$ is an output-stable pair, then by Theorem 2.2 (2) $\mathcal{G}_{C,\mathbf{A}}$ satisfies the Stein equation (2.15) and hence in particular

$$\mathcal{G}_{C,\mathbf{A}} - A_1^*\mathcal{G}_{C,\mathbf{A}}A_1 - \cdots - A_d^*\mathcal{G}_{C,\mathbf{A}}A_d \geq 0.$$

We now show that, for the abelianized case, the inequality in the reverse Stein inequality satisfied by the abelianized observability gramian $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ can be strict in the strong sense that the quantity $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} - A_1^*\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}A_1 - \cdots - A_d^*\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}A_d$ is not even positive semidefinite. As an example, let

$$C = \begin{bmatrix}1 & 0 & 0\end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A straightforward calculation shows that

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} - A_1^*\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}A_1 - \cdots - A_d^*\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}A_d = \begin{bmatrix} \frac{7}{8} & \frac{5}{8} & \frac{3}{8} \\ \frac{5}{8} & 0 & \frac{1}{4} \\ \frac{3}{8} & \frac{1}{4} & 0 \end{bmatrix},$$

which is not positive semidefinite.

Condition (3.12) is worth a formal definition.

Definition 3.5 *Let $C \in \mathcal{L}(\mathcal{X},\mathcal{Y})$. A d-tuple $\mathbf{A} = (A_1,\ldots,A_d)$ of bounded operators on $\mathcal{X}$ will be called* C-abelian *if* (3.12) *holds.*

One obvious way for a given operator d-tuple $\mathbf{A}$ to be C-abelian is for $\mathbf{A}$ itself to be commutative, i.e., for $A_iA_j = A_jA_i$ for all $1 \leq i,j \leq d$. We next show that, under an observability assumption, this is the only way.

Proposition 3.6 *Suppose that the output-stable pair $(C,\mathbf{A})$ is observable and that $\mathbf{A}$ is C-abelian. Then the d-tuple $\mathbf{A}$ is commutative.*

Proof Since $\mathbf{A}$ is C-abelian, relations (3.12) hold. Fix $i, j \in \{1, \ldots, d\}$ and note that by (3.12),

$$C\mathbf{A}^v A_i A_j = C\mathbf{A}^{vij} = C\mathbf{A}^{vji} = C\mathbf{A}^v A_j A_i \quad \text{for every } v \in \mathcal{F}_d,$$

since $\mathbf{a}(vij) = \mathbf{a}(vji)$. Thus,

$$C\mathbf{A}^v (A_i A_j - A_j A_i)x = 0$$

for every $v \in \mathcal{F}_d$ and $x \in \mathcal{X}$. Since the pair $(C, \mathbf{A})$ is observable, we have by (2.13)

$$(A_i A_j - A_j A_i)x = 0$$

holding for every $x \in \mathcal{X}$, which proves the commutativity relations

$$A_i A_j = A_j A_i \quad \text{for} \;\; i, j = 1, \ldots, d$$

and completes the proof. □

Corollary 3.7 *Suppose that $(C, \mathbf{A})$ is an observable output-stable pair. Then the abelianized observability gramian coincides with the observability gramian*

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}}$$

if and only if the operator d-tuple $\mathbf{A}$ is commutative.

Proof Combine (2) $\Longleftrightarrow$ (3) in Proposition 3.3 with Proposition 3.6. □

We next show that the observability gramian always dominates the abelianized observability gramian.

Proposition 3.8 *Let $(C, \mathbf{A})$ be an output-stable pair. Then:*

(1) *$(C, \mathbf{A})$ is also $\mathbf{a}$-output-stable with*

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} \le \mathcal{G}_{C,\mathbf{A}}. \tag{3.29}$$

(2) *Equality occurs in (3.29) if and only if $\mathbf{A}$ is C-abelian:*

$$C\mathbf{A}^v = C\mathbf{A}^u, \; \textit{whenever } u, v \in \mathcal{F}_d \textit{ with } \mathbf{a}(u) = \mathbf{a}(v).$$

Proof Note that the second statement in Proposition 3.8 is just a restatement of (2) $\Longleftrightarrow$ (3) in Proposition 3.3. Thus it suffices only to prove the first statement.

By definition, output-stability of $(C, \mathbf{A})$ simply means that $\mathcal{G}_{C,\mathbf{A}}$ is bounded, while $\mathbf{a}$-output stability means that $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ is bounded. The fact that $\mathbf{a}$-output stability follows from output-stability therefore follows immediately from the general inequality (3.29). Thus it suffices to prove (3.29). For this purpose, recall that

$$\langle \mathcal{G}_{C,\mathbf{A}} x, x \rangle = \left\| \mathcal{O}_{C,\mathbf{A}} x \right\|^2_{\ell^2_{\mathcal{Y}}(\mathcal{F}_d)} = \sum_{v \in \mathcal{F}_d} \| C\mathbf{A}^v x \|^2_{\mathcal{Y}}$$

while

$$\langle \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} x, x \rangle = \left\| \mathcal{O}^{\mathbf{a}}_{C,\mathbf{A}} x \right\|^2_{\ell^2_{\mathcal{Y}}(\mathbb{Z}^d_+)} = \sum_{\mathbf{n} \in \mathbb{Z}^d_+} \frac{\mathbf{n}!}{|\mathbf{n}|!} \left\| \sum_{v \in \mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^v x \right\|^2_{\mathcal{Y}}.$$

By the Cauchy–Schwarz inequality we have

$$\left\| \sum_{v\in\mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^v x \right\|_{\mathcal{Y}}^2 \le \left(\sum_{v\in\mathbf{a}^{-1}(\mathbf{n})} \|C\mathbf{A}^v x\|_{\mathcal{Y}} \right)^2 \le \sum_{v\in\mathbf{a}^{-1}(\mathbf{n})} \|C\mathbf{A}^v x\|_{\mathcal{Y}}^2 \cdot \frac{|\mathbf{n}|!}{\mathbf{n}!}.$$

Therefore,

$$\begin{aligned}\left\|\mathcal{O}^{\mathbf{a}}_{C,\mathbf{A}} x\right\|^2_{\ell^2_{w,\mathcal{Y}}(\mathbb{Z}^d_+)} &= \sum_{\mathbf{n}\in\mathbb{Z}^d_+} \frac{\mathbf{n}!}{|\mathbf{n}|!} \left\| \sum_{v\in\mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^v x \right\|^2_{\mathcal{Y}} \\ &\le \sum_{\mathbf{n}\in\mathbb{Z}^d_+} \sum_{v\in\mathbf{a}^{-1}(\mathbf{n})} \|C\mathbf{A}^v x\|^2_{\mathcal{Y}} \\ &= \sum_{v\in\mathcal{F}_d} \|C\mathbf{A}^v x\|^2_{\mathcal{Y}} = \|\mathcal{O}_{C,\mathbf{A}} x\|^2\end{aligned}$$

and (3.29) follows as wanted. □

Example 3.9 The converse of Proposition 3.8 part (1) can fail, i.e., *there exists an output pair* $(C, \mathbf{A})$ *which is* $\mathbf{a}$*-output-stable but not output-stable.* For example take

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then $C(I - \lambda_1 A_1 - \lambda_2 A_2)^{-1} = \begin{bmatrix} 1 & 2\lambda_1 & 2\lambda_2 \end{bmatrix}$. Hence

$$\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} \colon x \to C(I - \lambda_1 A_1 - \lambda_2 A_2)^{-1} x$$

maps $\mathcal{X} = \mathbb{C}^3$ into $\mathcal{H}(k_2)$ and thus $(C, \mathbf{A})$ is $\mathbf{a}$-output stable. To show that $(C, \mathbf{A})$ is not output stable, note that

$$(A_1 A_2)^n = \begin{bmatrix} 2^n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (-2)^n \end{bmatrix}$$

and therefore, $C(A_1 A_2)^n = \begin{bmatrix} 2^n & 0 & 0 \end{bmatrix}$, so that for $x = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\top$,

$$\sum_{v\in\mathcal{F}_d} \|C\mathbf{A}^v x\|^2_{\mathbb{C}} \ge \sum_{n\ge 0} 2^n = \infty$$

and therefore, the pair $(C, \mathbf{A})$ is not output-stable. We conclude that $\mathbf{a}$-output-stability has no obvious characterization in terms of positive semidefiniteness of some solution of a Stein inequality as in the noncommutative case (see Theorem 2.2 (2)).

As a corollary of the gramian inequality (3.29) in Proposition 3.8, we have the following.

Corollary 3.10 *Let* $(C, \mathbf{A})$ *be an output-stable pair. Then:*

(1) $\operatorname{Ker} \mathcal{G}_{C,\mathbf{A}} \subset \operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$. *Hence, if* $(C, \mathbf{A})$ *is* $\mathbf{a}$*-observable (respectively, exactly* $\mathbf{a}$*-observable), then* $(C, \mathbf{A})$ *is also observable (respectively, exactly observable).*
(2) *The subspace* $\operatorname{Ker} \mathcal{G}_{C,\mathbf{A}} = \operatorname{Ker} \mathcal{O}_{C,\mathbf{A}}$ *is invariant under the operator* A_j *for each* $j = 1, \ldots, d$.

(3) *The subspace* $\operatorname{Ker}\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ *is invariant under* A_j *for each* $j = 1,\dots,d$ *if and only if* $\operatorname{Ker}\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \operatorname{Ker}\mathcal{G}_{C,\mathbf{A}}$.

Proof Statement (1) is an immediate consequence of the inequality (3.29). Statement (2) is easily checked from the definition of $\mathcal{O}_{C,\mathbf{A}}$. Sufficiency in statement (3) is then a consequence of statement (2). It remains only to verify necessity in statement (3).

Assume therefore that $\operatorname{Ker}\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ is invariant under A_j for each $j = 1,\dots,d$. Let x be a vector in $\operatorname{Ker}\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$. Then by the assumed invariance, $\mathbf{A}^u x \in \operatorname{Ker}\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ for every $u \in \mathcal{F}_d$. Then we have

$$\operatorname{Ker} \sum_{v\in\mathbf{a}^{-1}(\mathbf{n})} C\mathbf{A}^{vu}x = 0 \quad \text{for every } \ u \in \mathcal{F}_d.$$

Then letting $\mathbf{n} = 0$ we get $C\mathbf{A}^u x = 0$ for every $u \in \mathcal{F}_d$ and therefore, $x \in \operatorname{Ker}\mathcal{G}_{C,\mathbf{A}}$. Thus, $\operatorname{Ker}\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} \subset \operatorname{Ker}\mathcal{G}_{C,\mathbf{A}}$ and since the reverse inclusion holds by the first statement, equality follows. □

Example 3.11 We observed in part (1) of Corollary 3.10 that **a**-observability for an output-stable pair $(C,\mathbf{A})$ implies observability. We now give an example to show that the converse can fail, i.e., there exists an output-stable observable pair which is not **a**-observable. For this purpose, let $d = 2$, $\mathcal{X} = \mathbb{C}^4$, $\mathcal{Y} = \mathbb{C}$, $C = \begin{bmatrix}0\ 0\ 0\ 1\end{bmatrix}$ and $\mathbf{A} = (A_1, A_2)$, where

$$A_1 = \begin{bmatrix} -\frac{1}{16} & \frac{1}{16} & 0 & 0 \\ -\frac{1}{16} & \frac{1}{16} & -\frac{1}{16} & \frac{1}{16} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{16} & \frac{1}{16} \end{bmatrix}, \qquad A_2 = \begin{bmatrix} \frac{1}{16} & 0 & 0 & -\frac{1}{16} \\ -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} \\ \frac{1}{16} & -\frac{1}{16} & \frac{1}{16} & -\frac{1}{16} \\ -\frac{1}{16} & 0 & 0 & -\frac{1}{16} \end{bmatrix}.$$

Then the pair $(C,\mathbf{A})$ is output stable. Now we show that $(C,\mathbf{A})$ is observable but not **a**-observable. Indeed, straightforward verifications give

$$CA_1 = \frac{1}{16}\begin{bmatrix}0\ 0\ {-1}\ 1\end{bmatrix}, \quad CA_2 = -\frac{1}{16}\begin{bmatrix}1\ 0\ 0\ 1\end{bmatrix},$$

and

$$CA_1A_2 = -\frac{1}{256}\begin{bmatrix}2\ {-1}\ 1\ 0\end{bmatrix}.$$

Now it is clear that $\operatorname{Ker} C \cap \operatorname{Ker} CA_1 \cap \operatorname{Ker} CA_2 \cap \operatorname{Ker} CA_1A_2 = 0$ which implies that $\bigcap_{v\in\mathcal{F}_d} \operatorname{Ker} C\mathbf{A}^v = 0$. Therefore, the pair $(C,\mathbf{A})$ is observable. To show that $(C,\mathbf{A})$ is not **a**-observable we first compute

$$I - \lambda_1 A_1 - \lambda_2 A_2 = \begin{bmatrix} 1 + \frac{\lambda_1}{16} - \frac{\lambda_2}{16} & -\frac{\lambda_1}{16} & 0 & \frac{\lambda_2}{16} \\ \frac{\lambda_1}{16} + \frac{\lambda_2}{16} & 1 - \frac{\lambda_1}{16} + \frac{\lambda_2}{16} & \frac{\lambda_1}{16} + \frac{\lambda_2}{16} & -\frac{\lambda_1}{16} + \frac{\lambda_2}{16} \\ -\frac{\lambda_2}{16} & \frac{\lambda_2}{16} & 1 - \frac{\lambda_2}{16} & \frac{\lambda_2}{16} \\ \frac{\lambda_2}{16} & 0 & \frac{\lambda_1}{16} & 1 - \frac{\lambda_1}{16} + \frac{\lambda_2}{16} \end{bmatrix}.$$

A straightforward calculation gives

$$\begin{aligned} d(\lambda_1,\lambda_2) &:= \det(I - \lambda_1 A_1 - \lambda_2 A_2) \\ &= 1 - \frac{\lambda_1}{16} + \frac{\lambda_1\lambda_2}{128} - \frac{\lambda_1^2\lambda_2}{2048} - \frac{\lambda_2^2}{64} + \frac{\lambda_1\lambda_2^2}{2048} - \frac{\lambda_1\lambda_2^3}{16384} + \frac{\lambda_2^4}{16384}. \end{aligned}$$

Note that

$$\begin{bmatrix} y_1 \; y_2 \; y_3 \; y_4 \end{bmatrix} := C(I - \lambda_1 A_1 - \lambda_2 A_2)^{-1}$$

is the bottom row of the matrix $(I - \lambda_1 A_1 + \lambda_2 A_2)^{-1}$ and we use the standard adjoint formula for the inverse of a matrix to get

$$y_2 = \frac{1}{d(\lambda_1, \lambda_2)} \cdot \begin{vmatrix} 1 + \frac{\lambda_1}{16} - \frac{\lambda_2}{16} & -\frac{\lambda_1}{16} & 0 \\ -\frac{\lambda_2}{16} & \frac{\lambda_2}{16} & 1 - \frac{\lambda_2}{16} \\ \frac{\lambda_2}{16} & 0 & \frac{\lambda_1}{16} \end{vmatrix} \equiv 0.$$

Then it follows that the nonzero vector $x = \begin{bmatrix} 0 \; 1 \; 0 \; 0 \end{bmatrix}^\top$ satisfies

$$C(I - \lambda_1 A_1 - \lambda_2 A_2)^{-1} x \equiv 0$$

and therefore, the pair $(C, \mathbf{A})$ is not $\mathbf{a}$-observable.

3.2 Observability-operator range spaces and reproducing kernel Hilbert spaces: the commutative-variable case

We seek the analogue of Theorem 1.2 for the commuting multivariable case. We extend multivariable power notation (3.1) to any d-tuple $\mathbf{A} = (A_1, \ldots, A_d)$ of commuting operators on a space $\mathcal{X}$:

$$\mathbf{A}^{\mathbf{n}} := A_1^{n_1} A_2^{n_2} \ldots A_d^{n_d}. \tag{3.30}$$

Note the connection between the commutative powers $\mathbf{A}^{\mathbf{n}}$ (with $\mathbf{n} \in \mathbb{Z}_+^d$) and the noncommutative powers $\mathbf{A}^v$ (with $v \in \mathcal{F}_d$) in case $\mathbf{A}$ is a commutative operator d-tuple:

$$\mathbf{A}^v = \mathbf{A}^{\mathbf{n}}, \quad \text{where} \quad \mathbf{n} = \mathbf{a}(v), \qquad \sum_{v \in \mathcal{F}_d \colon |v| = N} \mathbf{A}^{v^\top} X \mathbf{A}^v = \sum_{\mathbf{n} \in \mathbb{Z}_+^d \colon |\mathbf{n}| = N} \frac{N!}{\mathbf{n}!} \mathbf{A}^{*\mathbf{n}} X \mathbf{A}^{\mathbf{n}}$$

for any operator X on $\mathcal{X}$. In case $(C, \mathbf{A})$ is an output stable pair with $\mathbf{A}$ a commutative operator d-tuple, the formulas (3.7), (2.10), and (3.10) for $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$, $\mathcal{G}_{C,\mathbf{A}}$, and $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}}$ collapse (in view of (3.2)) to

$$\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} \colon \; x \mapsto C(I - Z(\boldsymbol{\lambda})A)^{-1} x = \sum_{\mathbf{n} \in \mathbb{Z}_+^d} \frac{|\mathbf{n}|!}{\mathbf{n}!} \left(C \mathbf{A}^{\mathbf{n}} x \right) \boldsymbol{\lambda}^{\mathbf{n}} \tag{3.31}$$

and

$$\mathcal{G}_{C,\mathbf{A}} = \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \sum_{\mathbf{n} \in \mathbb{Z}_+^d} \frac{|\mathbf{n}|!}{\mathbf{n}!} \mathbf{A}^{*\mathbf{n}} C^* C \mathbf{A}^{\mathbf{n}}. \tag{3.32}$$

We next observe that a natural commutative counterpart of operators S_j introduced in (2.25) are the operators M_{λ_j} of multiplication by the coordinate functions of $\mathbb{C}^d$ for $j = 1, \ldots, d$ acting as contractions on the Arveson space $\mathcal{H}_{\mathcal{Y}}(k_d)$. We will call the commuting d-tuple $\mathbf{M}_{\boldsymbol{\lambda}} := (M_{\lambda_1}, \ldots, M_{\lambda_d})$ the shift of $\mathcal{H}_{\mathcal{Y}}(k_d)$, whereas the commuting d-tuple $\mathbf{M}^*_{\boldsymbol{\lambda}} := (M^*_{\lambda_1}, \ldots, M^*_{\lambda_d})$ consisting of the adjoints of M_{λ_j}'s (in the metric of $\mathcal{H}_{\mathcal{Y}}(k_d)$) will be referred to as to the *backward shift*. Recall that monomials $\boldsymbol{\lambda}^{\mathbf{n}}$ form an orthogonal basis for $\mathcal{H}(k_d)$. As we have seen,

$$\langle \boldsymbol{\lambda}^{\mathbf{n}}, \boldsymbol{\lambda}^{\mathbf{m}} \rangle_{\mathcal{H}(k_d)} = \begin{cases} \dfrac{\mathbf{n}!}{|\mathbf{n}|!}, & \text{if} \quad \mathbf{n} = \mathbf{m}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.33}$$

A simple calculation based on (3.33) gives

$$M_{\lambda_j}^* \boldsymbol{\lambda}^{\mathbf{m}} = \frac{m_j}{|\mathbf{m}|} \boldsymbol{\lambda}^{\mathbf{m}-e_j} \ \ (m_j \geq 1) \quad \text{and} \quad M_{\lambda_j}^* \boldsymbol{\lambda}^{\mathbf{m}} = 0 \ \ (m_j = 0), \tag{3.34}$$

where $\mathbf{m} = (m_1, \ldots, m_d)$ and $e_j \in \mathbb{Z}_+^d$ is defined in (1.19). More generally,

$$\left(\mathbf{M}_{\boldsymbol{\lambda}}^*\right)^{\mathbf{n}} \ \boldsymbol{\lambda}^{\mathbf{m}} = \begin{cases} \dfrac{\mathbf{m}!|\mathbf{m}-\mathbf{n}|!}{|\mathbf{m}|!(\mathbf{m}-\mathbf{n})!} \boldsymbol{\lambda}^{\mathbf{m}-\mathbf{n}}, \text{ if } m_j \geq n_j \ \text{ for } j = 1, \ldots, d, \\ 0, \qquad\qquad\qquad\qquad \text{otherwise,} \end{cases} \tag{3.35}$$

where according to (3.30)

$$\left(\mathbf{M}_{\boldsymbol{\lambda}}^*\right)^{\mathbf{n}} := \left(M_{\lambda_1}^*\right)^{n_1} \left(M_{\lambda_2}^*\right)^{n_2} \cdots \left(M_{\lambda_d}^*\right)^{n_d}.$$

The following proposition includes the analogue of Proposition 2.9 for the present commutative setting.

Proposition 3.12 *Let $\mathbf{M}_{\boldsymbol{\lambda}}^*$ be the d-tuple of backward shifts on $\mathcal{H}_{\mathcal{Y}}(k_d)$ and let G: $\mathcal{H}_{\mathcal{Y}}(k_d) \to \mathcal{Y}$ be the operator of evaluation at $0 \in \mathbb{B}^d$*

$$G\colon \ f(\boldsymbol{\lambda}) \to f(0). \tag{3.36}$$

Then:

(1) *For every $f \in \mathcal{H}_{\mathcal{Y}}(k_d)$ and every $\boldsymbol{\lambda} \in \mathbb{B}^d$ we have*

$$f(\boldsymbol{\lambda}) - f(0) = \sum_{j=1}^{d} \lambda_j (M_{\lambda_j}^* f)(\boldsymbol{\lambda}). \tag{3.37}$$

(2) *The pair $(G, \mathbf{M}_{\boldsymbol{\lambda}}^*)$ is isometric:*

$$I - M_{\lambda_1} M_{\lambda_1}^* - \cdots - M_{\lambda_d} M_{\lambda_d}^* = G^* G. \tag{3.38}$$

(3) *The abelianized observability operator associated with the pair $(G, \mathbf{M}_{\boldsymbol{\lambda}}^*)$ is the identity operator:*

$$\widehat{\mathcal{O}}^{\mathbf{a}}_{G, \mathbf{M}_{\boldsymbol{\lambda}}^*} = I_{\mathcal{H}_{\mathcal{Y}}(k_d)}. \tag{3.39}$$

(4) *The d-tuple $\mathbf{M}_{\boldsymbol{\lambda}}^*$ is strongly stable, that is,*

$$\lim_{N \to \infty} \sum_{v \in \mathcal{F}_d \colon |v| = N} \|(\mathbf{M}_{\boldsymbol{\lambda}}^*)^v f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} = \lim_{N \to \infty} \sum_{\mathbf{n} \in \mathbb{Z}_+^d \colon |\mathbf{n}| = N} \frac{N!}{\mathbf{n}!} \|(\mathbf{M}_{\boldsymbol{\lambda}}^*)^{\mathbf{n}} f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} = 0 \tag{3.40}$$

for every $f \in \mathcal{H}_{\mathcal{Y}}(k_d)$.

Proof of (1) One can easily verify the identity (3.37) on monomials $y \cdot \boldsymbol{\lambda}^{\mathbf{m}}$ (with $y \in \mathcal{Y}$ and $\mathbf{m} \in \mathbb{Z}_+^d$) using (3.34). Then the result follows for all $f \in \mathcal{H}_{\mathcal{Y}}^2(k_d)$ by linearity and continuity.

Proof of (2) Note that G^*: $\mathcal{Y} \to \mathcal{H}_{\mathcal{Y}}(k_d)$ is the identification of a vector $y \in \mathcal{Y}$ with the constant function $y \in \mathcal{H}_{\mathcal{Y}}(k_d)$. We then see that (3.38) is simply the operator expression of (3.37).

Proof of (3) From (3.35) and (3.36) we see that

$$G\left(\mathbf{M}_{\boldsymbol{\lambda}}^{*}\right)^{\mathbf{n}} f=\frac{\mathbf{n}!}{|\mathbf{n}|!} f_{\mathbf{n}} \quad \text{if} \quad f(\boldsymbol{\lambda})=\sum_{\mathbf{m} \in \mathbb{Z}_{+}^{d}} f_{\mathbf{m}} \boldsymbol{\lambda}^{\mathbf{m}} \quad \text{and} \quad \mathbf{n} \in \mathbb{Z}_{+}^{d}$$

and therefore, according to definition (3.31),

$$\widehat{\mathcal{O}}_{G, \mathbf{M}_{\boldsymbol{\lambda}}^{*}}^{\mathbf{a}} f:=\sum_{\mathbf{n} \in \mathbb{Z}_{+}^{d}} \frac{|\mathbf{n}|!}{\mathbf{n}!}\left(G\left(\mathbf{M}_{\boldsymbol{\lambda}}^{*}\right)^{\mathbf{n}} f\right) \boldsymbol{\lambda}^{\mathbf{n}}=\sum_{\mathbf{n} \in \mathbb{Z}_{+}^{d}} f_{\mathbf{n}} \boldsymbol{\lambda}^{\mathbf{n}}=f(\boldsymbol{\lambda}).$$

Since the latter equality holds for every $f \in \mathcal{H}_{\mathcal{Y}}(k_d)$, (3.39) follows as asserted.

Proof of (4) This can be derived directly from (3.35) or via Proposition 2.7 since $\widehat{\mathcal{O}}_{G, \mathbf{M}_{\boldsymbol{\lambda}}^{*}}^{\mathbf{a}}=I$ and therefore, the pair $(G, \mathbf{M}_{\boldsymbol{\lambda}}^{*})$ is exactly observable. □

Remark 3.13 Note that in contrast to the noncommutative case (Proposition 2.9), the operator

$$R=\begin{bmatrix} M_{\lambda_1}^{*} \\ \vdots \\ M_{\lambda_d}^{*} \\ G \end{bmatrix}: \mathcal{H}_{\mathcal{Y}}(k_d) \rightarrow (\mathcal{H}_{\mathcal{Y}}(k_d))^{d} \oplus \mathcal{Y}$$

is not unitary (just isometric). A simple calculation shows that

$$I-RR^{*}=\begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix}: \begin{bmatrix} (\mathcal{H}_{\mathcal{Y}}(k_d))^{d} \\ \mathcal{Y} \end{bmatrix} \rightarrow \begin{bmatrix} (\mathcal{H}_{\mathcal{Y}}(k_d))^{d} \\ \mathcal{Y} \end{bmatrix},$$

where P is the orthogonal projection of $(\mathcal{H}_{\mathcal{Y}}(k_d))^d$ onto the subspace

$$\left\{h=\begin{bmatrix} h_1 \\ \vdots \\ h_d \end{bmatrix} \in (\mathcal{H}_{\mathcal{Y}}(k_d))^{d}: \sum_{j=1}^{d} \lambda_j h_j(\boldsymbol{\lambda}) \equiv 0\right\}.$$

If a pair $(C, \mathbf{A})$ is **a**-output stable, then the observability operator $\widehat{\mathcal{O}}_{C,\mathbf{A}}^{\mathbf{a}}:\ \mathcal{X} \rightarrow \mathcal{H}_{\mathcal{Y}}(k_d)$ is bounded and its range

$$\operatorname{Ran} \widehat{\mathcal{O}}_{C, \mathbf{A}}^{\mathbf{a}}:=\{C(I-Z(\boldsymbol{\lambda}) A)^{-1} x:\ x \in \mathcal{X}\} \tag{3.41}$$

is a linear manifold in $\mathcal{H}_{\mathcal{Y}}(k_d)$. We have the following partial analogues of part (3) of Theorem 2.8.

Theorem 3.14 *Let* $(C, \mathbf{A})$ *be an* **a**-*output stable pair. Then:*

(1) $\operatorname{Ran} \widehat{\mathcal{O}}_{C,\mathbf{A}}^{\mathbf{a}}$ *with the lifted norm*

$$\left\|C(I-Z(\boldsymbol{\lambda}) A)^{-1} x\right\|_{\mathcal{H}(K_{C, \mathbf{A}}^{\mathbf{a}})}=\|Q^{\mathbf{a}} x\|_{\mathcal{X}}, \tag{3.42}$$

where $Q^{\mathbf{a}}$ *is the orthogonal projection of* $\mathcal{X}$ *onto* $(\operatorname{Ker} \mathcal{G}_{C,\mathbf{A}}^{\mathbf{a}})^{\perp}$, *is isometrically equal to the reproducing kernel Hilbert space* $\mathcal{H}(K_{C,\mathbf{A}}^{\mathbf{a}})$ *with reproducing kernel* $K_{C,\mathbf{A}}^{\mathbf{a}}(\boldsymbol{\lambda}, \boldsymbol{\zeta})$ *given by*

$$K_{C, \mathbf{A}}^{\mathbf{a}}(\boldsymbol{\lambda}, \boldsymbol{\zeta})=C(I-Z(\boldsymbol{\lambda}) A)^{-1}(I-A^{*} Z(\boldsymbol{\zeta})^{*})^{-1} C^{*} \quad (\boldsymbol{\lambda}, \boldsymbol{\zeta} \in \mathbb{B}^{d}).$$

(2) $\overline{\operatorname{Ran}}\, \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ *with norm inherited from* $\mathcal{H}_{\mathcal{Y}}(k_d)$ *is a reproducing kernel Hilbert space* $\mathcal{H}(K^{\mathbf{a},-}_{C,\mathbf{A}})$ *with reproducing kernel* $K^{\mathbf{a},-}_{C,\mathbf{A}}(\boldsymbol{\lambda},\zeta)$ *given by*

$$K^{\mathbf{a},-}(\boldsymbol{\lambda},\boldsymbol{\zeta}) = C(I - Z(\boldsymbol{\lambda})A)^{-1}(\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{-1}(I - A^*Z(\boldsymbol{\zeta})^*)^{-1}C^* \quad (\boldsymbol{\lambda},\boldsymbol{\zeta} \in \mathbb{B}^d).$$

We next discuss separately the case where $\mathbf{A}$ is C-abelian and then the general case.

3.2.1 $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *for the case where* $\mathbf{A}$ *is C-abelian*

In case $(C,\mathbf{A})$ is an $\mathbf{a}$-output-stable pair with $\mathbf{A}$ C-abelian, then we have the following commutative analogue of Theorem 2.10.

Theorem 3.15 *Let* $(C,\mathbf{A})$ *be a contractive* $\mathbf{a}$*-output-stable pair such that operator d-tuple* $\mathbf{A}$ *is C-abelian. Then:*

(1) *The intertwining relations*

$$M^*_{\boldsymbol{\lambda}_j}\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} = \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}A_j \quad \text{for } j = 1,\ldots,d \tag{3.43}$$

hold, and hence the linear submanifold $\operatorname{Ran}\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ *of* $\mathcal{H}_{\mathcal{Y}}(k_d)$ *is* $\mathbf{M}^*_{\boldsymbol{\lambda}}$*-invariant.*

(2) *The operator* $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ *maps* $\mathcal{X}$ *contractively into* $\mathcal{H}_{\mathcal{Y}}(k_d)$*. This mapping is isometric if and only if* $(C,\mathbf{A})$ *is isometric and* $\mathbf{A}$ *is strongly stable.*

(3) *If* $\mathcal{M} := \operatorname{Ran}\mathcal{O}^{\mathbf{a}}_{C,\mathbf{A}}$ *is given the lifted norm* (3.42) *(so* $\mathcal{M}$ *is isometrically equal to* $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *by Theorem 3.14 (1)), then the difference-quotient inequality*

$$\sum_{j=1}^{d}\|M^*_{\lambda_j}f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} \le \|f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} - \|f(0)\|^2_{\mathcal{Y}}$$

holds for every $f \in \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$*. Moreover, the difference-quotient identity*

$$\sum_{j=1}^{d}\|M^*_{\lambda_j}f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} = \|f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} - \|f(0)\|^2_{\mathcal{Y}}$$

holds for every $f \in \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *if and only if the subspace* $(\operatorname{Ker}\mathcal{G}_{C,\mathbf{A}})^{\perp}$ *is* $\mathbf{A}$*-invariant and the restriction* $(C^0,\mathbf{A}^0)$ *(defined in* (2.53)*) of* $(C,\mathbf{A})$ *to the subspace* $(\operatorname{Ker}\mathcal{G}_{C,\mathbf{A}})^{\perp}$ *is isometric.*

Proof By (3.31) and (3.34), we have for every $x \in \mathcal{X}$,

$$
\begin{aligned}
(M_{\lambda_j})^* \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x &= (M_{\lambda_j})^* \left(\sum_{\mathbf{n} \in \mathbb{Z}^d_+} \frac{|\mathbf{n}|!}{\mathbf{n}!} C \mathbf{A}^{\mathbf{n}} x \cdot \boldsymbol{\lambda}^{\mathbf{n}} \right) \\
&= \sum_{\mathbf{n} \in \mathbb{Z}^d_+} \frac{|\mathbf{n}|!}{\mathbf{n}!} \frac{n_j}{|\mathbf{n}|} C \mathbf{A}^{\mathbf{n}} x \cdot \boldsymbol{\lambda}^{\mathbf{n}-e_j} \\
&= \sum_{\mathbf{n} \in \mathbb{Z}^d_+} \frac{|\mathbf{n}-e_j|!}{(\mathbf{n}-e_j)!} C \mathbf{A}^{\mathbf{n}-e_j} \cdot \boldsymbol{\lambda}^{\mathbf{n}-e_j} A_j x \\
&= \left(\sum_{\mathbf{n} \in \mathbb{Z}^d_+} \frac{|\mathbf{n}|!}{\mathbf{n}!} C \mathbf{A}^{\mathbf{n}} \boldsymbol{\lambda}^{\mathbf{n}} \right) \cdot A_j x \\
&= \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} \cdot A_j x
\end{aligned}
$$

and (3.43) follows. This completes the proof of statement (1) in the theorem.

Since the pair $(C, \mathbf{A})$ is contractive and $\mathbf{A}$ is C-abelian, we have

$$
\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}} \le Q = Q^{\mathbf{a}} \le I.
$$

Therefore,

$$
\|\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x\|_{\mathcal{H}_{\mathcal{Y}}(k_d)} = \langle \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} x, x \rangle^{\frac{1}{2}}_{\mathcal{X}} \le \|Q^{\mathbf{a}} x\|_{\mathcal{X}} = \|\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x\|_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} \le \|x\|_{\mathcal{X}}. \tag{3.44}
$$

Now the arguments used in the proof of Theorem 2.10 can be used to prove the remaining statements in the theorem. □

For the converse direction we have the following result.

Theorem 3.16 *Let $\mathcal{M}$ be a Hilbert space of $\mathcal{Y}$-valued functions included into $\mathcal{H}_{\mathcal{Y}}(k_d)$ and let us assume that $\mathcal{M}$ is $\mathbf{M}^*_{\boldsymbol{\lambda}}$-invariant.*

(1) *If the inequality*

$$
\sum_{j=1}^{d} \|M^*_{\lambda_j} f\|^2_{\mathcal{M}} \le \|f\|^2_{\mathcal{M}} - \|f(0)\|^2_{\mathcal{Y}} \tag{3.45}
$$

holds for every $f \in \mathcal{M}$, then $\mathcal{M} = \operatorname{Ran} \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ for a contractive and exactly observable (with respect to $\mathcal{M}$) pair $(C, \mathbf{A})$ with the commutative d-tuple $\mathbf{A} = (A_1, \ldots, A_d)$. In particular, $\mathcal{M}$ is contractively included in $\mathcal{H}_{\mathcal{Y}}(k_d)$.

(2) *If the equality*

$$
\sum_{j=1}^{d} \|M^*_{\lambda_j} f\|^2_{\mathcal{M}} = \|f\|^2_{\mathcal{M}} - \|f(0)\|^2_{\mathcal{Y}} \tag{3.46}
$$

holds for every $f \in \mathcal{M}$, then $\mathcal{M} = \operatorname{Ran} \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ for an isometric and exactly observable (with respect to $\mathcal{M}$) pair $(C, \mathbf{A})$ with the commutative d-tuple $\mathbf{A}$. By part (1), $\mathcal{M}$ is contractively included in $\mathcal{H}_{\mathcal{Y}}(k_d)$. Moreover, it is isometrically included in $\mathcal{H}_{\mathcal{Y}}(k_d)$ if and only if the restriction of the backward shift $\mathbf{M}^*_{\boldsymbol{\lambda}}$ to $\mathcal{M}$ is strongly stable, i.e.,

$$
\lim_{N \to \infty} \sum_{\mathbf{n} \in \mathbb{Z}^d_+ : \, |\mathbf{n}| = N} \frac{N!}{\mathbf{n}!} \|(\mathbf{M}^*_{\boldsymbol{\lambda}})^{\mathbf{n}} f\|^2_{\mathcal{M}} = 0 \quad \text{for every } f \in \mathcal{M}. \tag{3.47}
$$

Proof Define operators $C\colon \mathcal{M} \to \mathcal{Y}$ and $A_1, \ldots, A_d\colon \mathcal{M} \to \mathcal{M}$ by

$$C = G|_{\mathcal{M}}\colon\ f \to f(0) \quad \text{and} \quad A_j = M^*_{\lambda_j}|_{\mathcal{M}}\ \ (j = 1, \ldots, d). \tag{3.48}$$

Thus, the d-tuple $\mathbf{A}$ is the restriction of the backward-shift tuple $\mathbf{M}^*_{\boldsymbol{\lambda}}$ to $\mathcal{M}$. By part (3) of Proposition 3.12, it follows that $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} = I_{\mathcal{M}}$ and thus, the pair $(C, \mathbf{A})$ is exactly observable (with respect to $\mathcal{M}$) and the range of the associated observability operator $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ coincides (algebraically) with $\mathcal{M}$. Now we write (3.45) in terms of operators (3.48) as

$$\sum_{j=1}^{d} \|A_j f\|^2_{\mathcal{M}} \le \|f\|^2_{\mathcal{M}} - \|Cf\|^2_{\mathcal{Y}} \qquad (f \in \mathcal{M})$$

and conclude that the pair $(C, \mathbf{A})$ is contractive. Similarly, assumption (3.46) means that the chosen pair $(C, \mathbf{A})$ is isometric. Furthermore, if $\mathcal{M}$ is included in $\mathcal{H}_{\mathcal{Y}}(k_d)$ isometrically, relation (3.47) holds since $\mathbf{M}^*_{\boldsymbol{\lambda}}$ is strongly stable (see part (4) of Proposition 3.12). Conversely, if (3.47) holds, that is, if the commutative d-tuple $\mathbf{A} = (A_1, \ldots, A_d)$ is strongly stable on $\mathcal{M}$, the Stein equation (2.15) has a unique positive semidefinite solution. Since the pair $(C, \mathbf{A})$ is isometric (recall that we are proving isometrical inclusion under assumption (3.46)), this unique solution is the identity operator. On the other hand the observability gramian $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}}$ defined by the convergent series (3.32) satisfies the same Stein equation (as observed in part (2) of Theorem 2.2). Thus, $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = \mathcal{G}_{C,\mathbf{A}} = I$. Note that the inequality (2.18) holds with $H = \mathcal{G}_{C,\mathbf{A}} = I$, i.e.,

$$\sum_{v \in \mathcal{F}_d\colon |v| < N} \mathbf{A}^{*v^\top} C^* C A^v \le I - \sum_{v \in \mathcal{F}_d\colon |v| = N+1} \mathbf{A}^{*v^\top} \mathbf{A}^v.$$

Taking strong limits as $N \to \infty$ and noting that $I = \mathcal{G}_{C,\mathbf{A}} = \sum_{v \in \mathcal{F}_d} \mathbf{A}^{*v^\top} C^* C \mathbf{A}^v$ then gives

$$I \le I - \text{s-lim}_{N\to\infty} \sum_{v \in \mathcal{F}_d\colon |v| = N} \mathbf{A}^{*v^\top} \mathbf{A}^v$$

from which the strong-stability of $\mathbf{A}$ follows. Then $\mathcal{M} = \operatorname{Ran} \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ is isometrically included in $\mathcal{H}_{\mathcal{Y}}(k_d)$ by statement (2) in Theorem 3.15. □

We have the following analogue of Theorem 2.13 for the present commutative situation.

Theorem 3.17 *Suppose that* $(C, \mathbf{A})$ *and* $(\widetilde{C}, \widetilde{\mathbf{A}})$ *are two observable output-stable pairs with both* $\mathbf{A}$ *and* $\widetilde{\mathbf{A}}$ *commutative such that* $K^{\mathbf{a}}_{C,\mathbf{A}}(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = K^{\mathbf{a}}_{\widetilde{C},\widetilde{\mathbf{A}}}(\boldsymbol{\lambda}, \boldsymbol{\zeta})$ *for all* $\boldsymbol{\lambda}, \boldsymbol{\zeta} \in \mathbb{B}^d$. *Then there is a unitary operator* $U\colon \mathcal{X} \to \widetilde{\mathcal{X}}$ *such that*

$$C = \widetilde{C} U \quad \textit{and} \quad A_j = U^{-1} \widetilde{A}_j U \quad \textit{for}\ \ j = 1, \ldots, d. \tag{3.49}$$

Proof Suppose that $(C, \mathbf{A})$ and $(\widetilde{C}, \widetilde{\mathbf{A}})$ are as in the hypothesis of the theorem. The identity of the kernels $K^{\mathbf{a}}_{C,\mathbf{A}}$ and $K^{\mathbf{a}}_{\widetilde{C},\widetilde{\mathbf{A}}}$ implies equality of the respective coefficients of $\boldsymbol{\lambda}^{\mathbf{n}} \boldsymbol{\zeta}^{\mathbf{m}}$ for each $\mathbf{n}, \mathbf{m} \in \mathbb{Z}^d_+$:

$$\frac{|\mathbf{n}|!}{\mathbf{n}!} \frac{|\mathbf{m}|!}{\mathbf{m}!} C \mathbf{A}^{\mathbf{n}} \mathbf{A}^{*\mathbf{m}} C^* = \frac{|\mathbf{n}|!}{\mathbf{n}!} \frac{|\mathbf{m}|!}{\mathbf{m}!} \widetilde{C} \widetilde{\mathbf{A}}^{\mathbf{n}} \widetilde{\mathbf{A}}^{*\mathbf{m}} \widetilde{C}^*.$$

If we define a mapping U by

$$U\colon \mathbf{A}^{*\mathbf{n}}C^*y \mapsto \widetilde{\mathbf{A}}^{*\mathbf{n}}\widetilde{C}^*y, \tag{3.50}$$

it follows that U extends by linearity to an isometry from

$$\mathcal{D} := \operatorname{span}\{\mathbf{A}^{*\mathbf{m}}C^*y\colon \mathbf{m} \in \mathbb{Z}_+^d \text{ and } y \in \mathcal{Y}\}$$

onto

$$\mathcal{R} := \operatorname{span}\{\widetilde{\mathbf{A}}^{*\mathbf{m}}\widetilde{C}^*y\colon \mathbf{m} \in \mathbb{Z}_+^d \text{ and } y \in \mathcal{Y}\}$$

Since both $(C, \mathbf{A})$ and $(\widetilde{C}, \widetilde{\mathbf{A}})$ are observable, we see that $\mathcal{D}$ is dense in $\mathcal{X}$ and that $\mathcal{R}$ is dense in $\widetilde{\mathcal{X}}$. Hence U extends to a unitary operator from $\mathcal{X}$ onto $\widetilde{\mathcal{X}}$ by continuity. From the defining equations (3.50) for U we see that

$$UC^* = \widetilde{C}^* \quad \text{and} \quad UA_j^* = \widetilde{A}_j^*U.$$

By taking adjoints and using that U is unitary, we arrive at the intertwining equations (3.49) as wanted. □

Theorem 3.17 can be adapted to give the following result concerning containment between two backward-shift-invariant subspaces rather than equality; the finite-dimensional case appears as Proposition 1.2 in Bolotnikov and Rodman (2002).

Theorem 3.18 *Let $\mathcal{M}$ and $\widetilde{\mathcal{M}}$ be two backward-shift-invariant subspaces of the Arveson space $\mathcal{H}_{\mathcal{Y}}(k_d)$ with realizations*

$$\mathcal{M} = \operatorname{Ran}\widehat{\mathcal{O}}_{C,\mathbf{A}} \quad \textit{and} \quad \widetilde{\mathcal{M}} = \operatorname{Ran}\widehat{\mathcal{O}}_{\widetilde{C},\widetilde{\mathbf{A}}}, \tag{3.51}$$

where the d-tuples $\mathbf{A} = (A_1, \ldots, A_d) \in \mathcal{X}^d$ and $\widetilde{\mathbf{A}} = (\widetilde{A}_1, \ldots, \widetilde{A}_d) \in \widetilde{\mathcal{X}}^d$ are commutative and strongly stable and the pairs $(C, \mathbf{A})$ and $(\widetilde{C}, \widetilde{\mathbf{A}})$ are isometric. Then $\mathcal{M} \subseteq \widetilde{\mathcal{M}}$ if and only if there exists an isometry $V\colon \mathcal{X} \to \widetilde{\mathcal{X}}$ such that

$$C = V\widetilde{C} \quad \textit{and} \quad VA_j = \widetilde{A}_jV \ \ (j = 1, \ldots, d). \tag{3.52}$$

Proof The necessity part is clear. For the sufficiency part, assume that $\mathcal{M} \subseteq \widetilde{\mathcal{M}}$. By Theorem 3.15, there exist unitary operators $U\colon \mathcal{M} \to \mathcal{X}$ and $\widetilde{U}\colon \widetilde{\mathcal{M}} \to \widetilde{\mathcal{X}}$ such that

$$U^*A_jU = M_{\lambda_j}^*|_{\mathcal{M}}, \quad \widetilde{U}^*\widetilde{A}_j\widetilde{U} = M_{\lambda_j}^*|_{\mathcal{M}} \quad (j = 1, \ldots, d)$$

and

$$CU = G|_{\mathcal{M}}, \quad \widetilde{C}\widetilde{U} = G|_{\mathcal{M}},$$

where the operator $G : \mathcal{H}_{\mathcal{Y}} \to \mathcal{Y}$ is defined in (3.36). Let $\mathcal{I}\colon \mathcal{M} \to \widetilde{\mathcal{M}}$ be the inclusion operator. Clearly $\mathcal{I}$ is isometric. Then the operator $V = U^*\mathcal{I}\widetilde{U}\colon \mathcal{X} \to \widetilde{\mathcal{X}}$ is isometric and satisfies (3.52). □

3.2.2 $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$: The general case

In case the $\mathbf{a}$-output-stable pair $(C, \mathbf{A})$ is such that $\mathbf{A}$ is not C-abelian, it can happen that the associated reproducing kernel Hilbert space is not invariant under the backward-shift tuple $\mathbf{M}_{\lambda}^*$, as the following example shows.

Example 3.19 Let

$$C = \begin{bmatrix} \frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}.$$

Then a straightforward calculation gives

$$\begin{aligned} K^{\mathbf{a}}_{C,\mathbf{A}}(\boldsymbol{\lambda}, \boldsymbol{\zeta}) &= C(I - Z(\boldsymbol{\lambda})A)^{-1}(I - A^*Z(\boldsymbol{\zeta})^*)^{-1}C^* \\ &= \frac{3}{(4-\lambda_1\lambda_2)(4-\overline{\zeta}_1\overline{\zeta}_2)} \begin{bmatrix} 2 & \lambda_2 \\ \lambda_1 & 2 \end{bmatrix} \begin{bmatrix} 2 & \overline{\zeta}_1 \\ \overline{\zeta}_2 & 2 \end{bmatrix}. \end{aligned}$$

Thus, $K^{\mathbf{a}}_{C,\mathbf{A}}(\lambda, w)$ is positive definite on $\mathbb{B}^2 \times \mathbb{B}^2$ and the space $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ is spanned by the two rational functions

$$f_1(\lambda) = \frac{4}{4-\lambda_1\lambda_2} \begin{bmatrix} 2 \\ \lambda_1 \end{bmatrix} \quad \text{and} \quad f_2(\lambda) = \frac{4}{4-\lambda_1\lambda_2} \begin{bmatrix} \lambda_2 \\ 2 \end{bmatrix}.$$

Furthermore, since

$$M^*_{\lambda_1}(\lambda_1^{n_1}\lambda_2^{n_2}) = \frac{n_1}{n_1+n_2}\lambda_1^{n_1-1}\lambda_2^{n_2}$$

and since

$$\frac{4\lambda_1}{4-\lambda_1\lambda_2} = \sum_{j=0}^{\infty} \frac{\lambda_1^{j+1}\lambda_2^j}{4^j}$$

it holds that

$$M^*_{\lambda_1}\left(\frac{4\lambda_1}{4-\lambda_1\lambda_2}\right) = \sum_{j=0}^{\infty} \frac{j+1}{2j+1}\left(\frac{\lambda_1\lambda_2}{4}\right)^j.$$

The latter function is rational if and only if the single-variable function $F(z) = \sum_{j=0}^{\infty} \frac{j+1}{2j+1} z^j$ is rational. By the well-known Kronecker theorem, F in turn is rational if and only if the associated infinite Hankel matrix

$$\mathbb{H} = [s_{i+j}]_{i,j=0}^{\infty}, \quad \text{where} \quad s_k = \frac{k+1}{2k+1}$$

has finite rank. However one can check that the finite Hankel matrices $\mathbb{H}_n = [s_{i+j}]_{i,j=0}^{n}$ have full rank for all $n = 0, 1, 2, \ldots$ and hence $F(z)$ is not rational. Therefore, $M^*_{\lambda_1} f_1$ does not belong to $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ and hence $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ is not invariant under $M^*_{\lambda_1}$.

For the general case, there is a simple replacement for $\mathbf{M}^*_{\boldsymbol{\lambda}}|_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})}$. Specifically, given an $\mathbf{a}$-output-stable pair $(C, \mathbf{A})$, we define an operator-tuple $\mathbf{T} = (T_1, \ldots, T_d)$ on $\operatorname{Ran} \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ by

$$T_j \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x = \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} A_j x \quad \text{for } x \in (\operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp} \text{ and } j = 1, \ldots, d. \tag{3.53}$$

We then have

$$\begin{aligned} f(\boldsymbol{\lambda}) - f(0) &= C(I - Z(\boldsymbol{\lambda})A)^{-1}x - Cx \\ &= C(I - Z(\boldsymbol{\lambda})A)^{-1}Z(\boldsymbol{\lambda})Ax \\ &= \sum_{j=1}^{d} \lambda_j C(I - Z(\boldsymbol{\lambda})A)^{-1}A_j x \\ &= \sum_{j=1}^{d} \lambda_j \cdot (\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} A_j x)(\boldsymbol{\lambda}) \\ &= \sum_{j=1}^{d} \lambda_j \cdot (T_j \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x)(\boldsymbol{\lambda}) = \sum_{j=1}^{d} \lambda_j \cdot (T_j f)(\boldsymbol{\lambda}). \end{aligned} \tag{3.54}$$

We next give the following analogue of Theorem 3.15 for the general case.

Theorem 3.20 *Let* $(C, \mathbf{A})$ *be a contractive pair with* $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ *and* $\mathbf{A} = (A_1, \dots, A_d) \in \mathcal{L}(\mathcal{X})^d$. *Then:*

(1) *The Z-transformed observability operator* $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ *is a contraction of* $\mathcal{X}$ *into the reproducing kernel Hilbert space* $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$. *It is an isometry if and only if the the pair* $(C, \mathbf{A})$ *is* $\mathbf{a}$*-observable.*

(2) *The space* $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *is contractively included in the Arveson space* $\mathcal{H}_{\mathcal{Y}}(k_d)$; *it is isometrically included in* $\mathcal{H}_{\mathcal{Y}}(k_d)$ *if and only if* $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ *(as an operator from* $\mathcal{X}$ *into* $\mathcal{H}_{\mathcal{Y}}(k_d)$*) is a partial isometry.*

(3) *For every function* $f \in \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *it holds that*

$$f(\boldsymbol{\lambda}) - f(0) = \sum_{j=1}^{d} \lambda_j (T_j f)(\boldsymbol{\lambda}) \quad (\boldsymbol{\lambda} \in \mathbb{B}^d) \tag{3.55}$$

and

$$\sum_{j=1}^{d} \|T_j f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} \leq \|f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} - \|f(0)\|^2_{\mathcal{Y}} \tag{3.56}$$

where $T_1, \dots, T_d \in \mathcal{L}(\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}}))$ *are the operators defined in* (3.53).

(4) *Equality holds in* (3.56) *for every* $f \in \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *if and only if the subspace* $(\mathrm{Ker}\, \mathcal{G}_{C,\mathbf{A}})^{\perp}$ *is* $\mathbf{A}$*-invariant and the restriction* $(C^0, \mathbf{A}^0)$ *(defined in* (2.53)*) of* $(C, \mathbf{A})$ *to the subspace* $(\mathrm{Ker}\, \mathcal{G}_{C,\mathbf{A}})^{\perp}$ *is isometric.*

(5) *If* $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *is isometrically included in* $\mathcal{H}_{\mathcal{Y}}(k_d)$, *then* $T_j = M^*_{\lambda_j} \setminus \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ *for* $j = 1, \dots, d$ *and therefore,* $\mathcal{H}_{\mathcal{Y}}(k_d)$ *is* $\mathbf{M}^*_{\boldsymbol{\lambda}}$*-invariant.*

Proof Since the pair $(C, \mathbf{A})$ is contractive, the identity operator $H = I_{\mathcal{X}}$ solves the Stein inequality (2.14). Then $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} \leq \mathcal{G}_{C,\mathbf{A}} \leq I_{\mathcal{X}}$ (by part (1) of Proposition 3.8 and part (2) of Theorem 2.2). Thus,

$$\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} \leq Q^{\mathbf{a}} \leq I_{\mathcal{X}},$$

where $Q^{\mathbf{a}}$ is the orthogonal projection of $\mathcal{X}$ onto $(\mathrm{Ker}\, \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp}$. Therefore it holds for every $x \in \mathcal{X}$ that

$$\|\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x\|_{\mathcal{H}_{\mathcal{Y}}(k_d)} = \langle \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} x, x \rangle^{\frac{1}{2}}_{\mathcal{X}} \leq \|Q^{\mathbf{a}} x\|_{\mathcal{X}} = \|\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x\|_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} \leq \|x\|_{\mathcal{X}}. \tag{3.57}$$

We have the equality instead of the first inequality in (3.57) if and only if $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} = Q^{\mathbf{a}}$, that is, if and only if $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ is a partial isometry. Furthermore, the second inequality in (3.57) can be replaced by equality if and only if $Q^{\mathbf{a}} = I_{\mathcal{X}}$, i.e., if and only if the pair $(C, \mathbf{A})$ is **a**-observable. This completes the proof of the two first assertions in the theorem. The multivariable difference-quotient relation (3.55) follows by the calculation (3.54). Furthermore, for every $x \in (\operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp}$,

$$\|\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x\|_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} = \|Q^{\mathbf{a}} x\|_{\mathcal{X}} = \|x\|_{\mathcal{X}}$$

and thus, $\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ maps unitarily $(\operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp}$ onto $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$. Therefore, by (3.53), T_j is unitarily equivalent to the compression of A_j to $(\operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp}$ and hence

$$\|T_j\| \le \|A_j\| \quad \text{for } j = 1, \ldots, d.$$

In particular, $T_j \in \mathcal{L}(\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}}))$. For an element $f = \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x \in \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$, we have

$$\begin{aligned}
\sum_{j=1}^{d} \|T_j f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} &= \sum_{j=1}^{d} \|T_j \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} x\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} \\
&= \sum_{j=1}^{d} \|\widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}} A_j x\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} \\
&= \sum_{j=1}^{d} \|Q^{\mathbf{a}} A_j x\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} \\
&\le \sum_{j=1}^{d} \|A_j x\|^2_{\mathcal{X}} \\
&\le \|x\|^2_{\mathcal{X}} - \|Cx\|^2_{\mathcal{Y}} = \|f\|^2_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} - \|f(0)\|^2_{\mathcal{Y}},
\end{aligned}$$

where the first inequality holds since $Q^{\mathbf{a}} \le I$ and the second since $(C, \mathbf{A})$ is a contractive pair. This proves inequality (3.56) and it is readily seen that equalities hold throughout in the last calculation for every $x \in (\operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp}$ if and only the subspace $(\operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp}$ is $\mathbf{A}$-invariant and the restriction $(C^0, \mathbf{A}^0)$ (defined in (2.53)) of $(C, \mathbf{A})$ to the subspace $(\operatorname{Ker} \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}})^{\perp}$ is isometric.

Finally, suppose that $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ is included isometrically in $\mathcal{H}_{\mathcal{Y}}(k_d)$. Then the assumption (3.56) becomes

$$\sum_{j=1}^{d} \|T_j f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} \le \|f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} - \|f(0)\|^2_{\mathcal{Y}} \text{ for every } f \in \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}}). \tag{3.58}$$

Then we take the inner product of both parts in equality (3.54) with f:

$$\langle f - f(0), f \rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)} = \|f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} - \|f(0)\|^2_{\mathcal{Y}}$$

and

$$\sum_{j=1}^{d} \langle M_{\lambda_j} T_j f, f \rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)} = \sum_{j=1}^{d} \langle T_j f, M^*_{\lambda_j} f \rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)}.$$

Thus,

$$\|f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} - \|f(0)\|^2_{\mathcal{Y}} = \sum_{j=1}^{d} \langle T_j f, M^*_{\lambda_j} f\rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)}. \tag{3.59}$$

For any f in $\mathcal{H}_{\mathcal{Y}}(k_d)$, applying the identity (3.38) to f and then taking the inner product with f gives us

$$\|f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} - \|f(0)\|^2_{\mathcal{Y}} = \sum_{j=1}^{d} \langle M_{\lambda_j} M^*_{\lambda_j} f, f\rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)} = \sum_{j=1}^{d} \|M^*_{\lambda_j} f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)}. \tag{3.60}$$

Now we conclude from (3.60), (3.59) and (3.58) that

$$\begin{aligned}\|f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} - \|f(0)\|^2_{\mathcal{Y}} &= \sum_{j=1}^{d} \|M^*_{\lambda_j} f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} \\ &= \sum_{j=1}^{d} \langle T_j f, M^*_{\lambda_j} f\rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)} \ge \sum_{j=1}^{d} \|T_j f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)}\end{aligned}$$

from which we get

$$\begin{aligned}0 &= \sum_{j=1}^{d} \|M^*_{\lambda_j} f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)} - \sum_{j=1}^{d} \langle T_j f, M^*_{\lambda_j} f\rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)}, \\ 0 &\ge -\sum_{j=1}^{d} \langle T_j f, M^*_{\lambda_j} f\rangle_{\mathcal{H}_{\mathcal{Y}}(k_d)} + \sum_{j=1}^{d} \|T_j f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)}.\end{aligned}$$

Adding these inequalities and using that $\sum_{j=1}^{d} \langle T_j f, M^*_{\lambda_j} f\rangle$ is real then gives

$$\begin{aligned}0 &\ge \sum_{j=1}^{d} \left(\|M^*_{\lambda_j} f\|^2 - \langle T_j f, M^*_{\lambda_j} f\rangle - \langle M^*_{\lambda_j} f, T_j f\rangle + \|T_j f\|^2 \right) \\ &= \sum_{j=1}^{d} \|M^*_{\lambda_j} f - T_j f\|^2_{\mathcal{H}_{\mathcal{Y}}(k_d)}.\end{aligned}$$

Therefore, $M^*_{\lambda_j} f = T_j f$ for $j = 1, \ldots, d$ and for every $f \in \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ as asserted. This completes the proof of Theorem 3.20. □

3.3 The Gleason problem: a uniqueness result

Let $\mathcal{M}$ be a Hilbert space of $\mathcal{Y}$-valued functions. A tuple $\mathbf{T} = (T_1, \ldots, T_d)$ of operators $T_j \in \mathcal{M}$ is called a solution of the *Gleason problem* (Gleason, 1964; Henkin, 1971) if relation (3.55) holds for every $f \in \mathcal{M}$. Let us say that $\mathbf{T}$ *is a contractive solution of the Gleason problem* if in addition

$$\sum_{j=1}^{d} \|T_j f\|^2_{\mathcal{M}} \le \|f\|^2_{\mathcal{M}} - \|f(0)\|^2_{\mathcal{Y}} \quad \text{for every} \quad f \in \mathcal{M} \tag{3.61}$$

or, equivalently, if the pair $(\mathbf{T}, G)$ is contractive where $G : \mathcal{M} \to \mathcal{Y}$ is defined by

$$G\colon\ f(\boldsymbol{\lambda}) \to f(0). \tag{3.62}$$

We have the following analogue of Theorem 3.16 characterizing contractively included subspaces $\mathcal{M}$ of $\mathcal{H}_{\mathcal{Y}}(k_d)$ of the form $\mathcal{M} = \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$; for the general case where $\mathcal{M}$ is not $\mathbf{M}^*_{\boldsymbol{\lambda}}$-invariant, one simply replaces $\mathbf{M}^*_{\boldsymbol{\lambda}}$ with some contractive solution $\mathbf{T}$ of the Gleason problem on $\mathcal{M}$.

Theorem 3.21 *Let $\mathcal{M}$ be a Hilbert space of $\mathcal{Y}$-valued functions and let us assume that there exists a contractive solution $\mathbf{T} = (T_1, \ldots, T_d)$ of the Gleason problem (i.e., $T_j \in \mathcal{L}(\mathcal{M})$ such that* (3.55) *and* (3.61) *hold for every $f \in \mathcal{M}$). Then $\mathcal{M}$ is isometrically equal to a reproducing kernel Hilbert space $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ for a contractive pair $(C, \mathbf{A})$. Therefore, $\mathcal{M}$ is contractively included in the Arveson space $\mathcal{H}_{\mathcal{Y}}(k_d)$.*

Proof Take $C = G|_{\mathcal{M}}$ where G is given by (3.62), $\mathbf{A} = \mathbf{T}$ on $\mathcal{M}$. Then (3.61) says that $(C, \mathbf{A})$ is contractive. Iteration of (3.55) says that, for each $f \in \mathcal{M}$,

$$f(\boldsymbol{\lambda}) = \sum_{j_1=1}^{d} \lambda_{j_1} \left[(T_{j_1} f)(0) + \sum_{j_2=1}^{d} \lambda_{j_2} \left[(T_{j_2} T_{j_1} f)(0) + \sum_{j_3=1}^{d} \lambda_{j_3} \left[(T_{j_3} T_{j_2} T_{j_1} f(0) + \right.\right.\right.$$
$$\left.\left.\left. \cdots + \sum_{j_k=1}^{d} \lambda_{j_k} \left[(T_{j_k} \cdots T_{j_2} T_{j_1} f)(0) + \cdots \right] \cdots \right] \right]\right].$$

This unravels to the tautology

$$f(\boldsymbol{\lambda}) = C(I - Z(\boldsymbol{\lambda})A)^{-1} f$$

so we recover $\mathcal{M}$ as $\mathcal{M} = \operatorname{Ran} \widehat{\mathcal{O}}^{\mathbf{a}}_{C,\mathbf{A}}$ with $\|C(I - Z(\cdot)A)^{-1} f\|_{\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})} = \|f\|_{\mathcal{M}}$, i.e., $\mathcal{M} = \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ isometrically. From the fact that $(C, \mathbf{A})$ is contractive, we have seen that $\mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} \le \mathcal{G}_{C,\mathbf{A}} \le I_{\mathcal{M}}$. Then

$$\|f\|^2_{\mathcal{H}^2_{\mathcal{Y}}(k_d)} = \|C(I - Z(\cdot)A)^{-1} f\|^2_{\mathcal{H}^2_{\mathcal{Y}}(k_d)} = \langle \mathcal{G}^{\mathbf{a}}_{C,\mathbf{A}} f, f \rangle_{\mathcal{M}} \le \|f\|^2_{\mathcal{M}}$$

and we also have the contractive inclusion property. □

Combining Theorems 3.20 and 3.21 gives the following uniqueness result for contractive solutions of the Gleason problem on a subspace $\mathcal{M}$ contained in $\mathcal{H}_{\mathcal{Y}}(k_d)$ isometrically.

Theorem 3.22 *Suppose that $\mathcal{M}$ is a subspace of $\mathcal{Y}$-valued functions contained in $\mathcal{H}_{\mathcal{Y}}(k_d)$ isometrically and that $\mathbf{T} = (T_1, \ldots, T_d)$ is a contractive solution of the Gleason problem on $\mathcal{M}$. Then $\mathcal{M}$ is $\mathbf{M}^*_{\boldsymbol{\lambda}}$-invariant and $\mathbf{T} = \mathbf{M}^*_{\boldsymbol{\lambda}}$.*

Proof By Theorem 3.21, there is a contractive pair $(C, \mathbf{A})$ so that $\mathcal{M} = \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ isometrically. As $\mathcal{M}$ is contained in $\mathcal{H}_{\mathcal{Y}}(k_d)$ isometrically, we conclude that $\mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ is contained in $\mathcal{H}_{\mathcal{Y}}(k_d)$ isometrically. Part (5) in Theorem 3.20 then asserts that the subspace $\mathcal{M} = \mathcal{H}(K^{\mathbf{a}}_{C,\mathbf{A}})$ is $\mathbf{M}^*_{\boldsymbol{\lambda}}$-invariant and that $T_j = M^*_{\lambda_j}$ for $j = 1, \ldots, d$. □

We note that the proof of Theorem 2.13 is like the proof of the State-Space-Isomorphism Theorem for structured noncommutative multidimensional linear systems in Ball et al. (2006). It is known that the State-Space-Isomorphism Theorem (and related Kalman reduction procedure) fails in general for commutative multidimensional linear systems—see, e.g., Galkowski (2005) for a recent account of the situation. The fact that uniqueness does hold in the special commutative situation in Theorem 3.17 shows that the technique in the proof of the State-Space-Isomorphism Theorem is salvageable in special commutative situations.

A uniqueness result for solutions of the Gleason problem somewhat different from that in Theorem 3.22 was obtained in Alpay and Dubi (2005); rather than assuming that $\mathbf{T}$ is a contractive solution of the Gleason problem on $\mathcal{M} = \mathcal{H}(K_{C,\mathbf{A}})$ contained isometrically in $\mathcal{H}_{\mathcal{Y}}(k_d)$ as in Theorem 3.22, Alpay and Dubi (2005) assume instead that $\mathbf{T}$ is a commutative solution of the Gleason problem and are then able to conclude that necessarily $\mathbf{T} = \mathbf{M}_{\boldsymbol{\lambda}}^*|_{\mathcal{M}}$. This latter result can be seen as an immediate consequence of our Theorem 3.17 above since, by the construction in the proof of Theorem 3.21, solutions $(C, \mathbf{A})$ of $K^{\mathbf{a}}_{C,\mathbf{A}} = K$ are in one-to-one correspondence with solutions $\mathbf{T}$ of the Gleason problem. We illustrate the preceding analysis by two examples.

Example 3.23 Consider the subspace $\mathcal{M} = \operatorname{span}\{1, \lambda_1, \lambda_2\} \subset \mathcal{H}(k_2)$ and define the operators $T_{a,1}$ and $T_{a,2}$ on $\mathcal{M}$ by

$$T_{a,1} : f \mapsto \beta + a\alpha\lambda_2, \quad T_{a,2} : f \mapsto \gamma - a\alpha\lambda_1 \tag{3.63}$$

where $f(\boldsymbol{\lambda}) = \alpha + \beta\lambda_1 + \gamma\lambda_2$ is the generic element in $\mathcal{M}$ and where a is a fixed complex number. It is readily checked that

$$f(\boldsymbol{\lambda}) - f(0) = \beta\lambda_1 + \gamma\lambda_2 = \lambda_1(T_{a,1}f)(\boldsymbol{\lambda}) + \lambda_2(T_{a,2}f)(\boldsymbol{\lambda})$$

so the tuple $(T_{a,1}, T_{a,2})$ solves the Gleason problem on $\mathcal{M}$. Let $A_{a,1}$ and $A_{a,2}$ be the matrices of $T_{a,1}$ and $T_{a,2}$ with respect to the basis $\{1, \lambda_1, \lambda_2\}$ of $\mathcal{M}$ and let C be the matrix of the operator $G : \mathcal{M} \to \mathcal{Y}$ defined in (3.62):

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad A_{a,1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ a & 0 & 0 \end{bmatrix}, \quad A_{a,2} = \begin{bmatrix} 0 & 0 & 1 \\ -a & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.64}$$

A straightforward calculation shows that

$$C(I - \lambda_1 A_{a,1} - \lambda_2 A_{a,2})^{-1} = \begin{bmatrix} 1 & \lambda_1 & \lambda_2 \end{bmatrix},$$

which realizes $\mathcal{M}$ as the range of the observability operator of a pair $(C, \mathbf{A}_a)$. Different choices of a in (3.64) lead to nonequivalent realizations of $\mathcal{M}$. Note that $A_{a,1}$ and $A_{a,2}$ do not commute unless $a = 0$, in which case the operators $T_{0,1}$ and $T_{0,2}$ are equal to backward shifts $M^*_{\lambda_1}$ and $M^*_{\lambda_2}$, respectively; in other words, the matrices

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad A_{0,1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_{0,2} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.65}$$

provide a commutative realization of $\mathcal{M}$ which is unique (up to unitary equivalence) by Theorem 3.17. Note also that the tuple $(T_{a,1}, T_{a,2})$ defined in (3.63) is never a contractive solution of the Gleason problem unless $a = 0$.

Example 3.24 Consider the subspace

$$\mathcal{M} = \operatorname{span}\left\{ \frac{4}{4 - \lambda_1\lambda_2}, \frac{\lambda_1}{4 - \lambda_1\lambda_2}, \frac{\lambda_2}{4 - \lambda_1\lambda_2} \right\} \subset \mathcal{H}(k_2)$$

and define the operators $T_{a,1}$ and $T_{a,2}$ on $\mathcal{M}$ by

$$T_{a,1} : f \mapsto \frac{\beta + a\alpha\lambda_2}{4-\lambda_1\lambda_2}, \quad T_{a,2} : f \mapsto \frac{\gamma + (1-a)\alpha\lambda_1}{4-\lambda_1\lambda_2},$$

where a s a fixed complex number and where

$$f(\boldsymbol{\lambda}) = \frac{4\alpha + \beta\lambda_1 + \gamma\lambda_2}{4-\lambda_1\lambda_2}$$

is the generic element in $\mathcal{M}$. Thus, $f(0) = \alpha$ and it is readily checked that

$$f(\boldsymbol{\lambda}) - f(0) = \frac{\alpha\lambda_1\lambda_2 + \beta\lambda_1 + \gamma\lambda_2}{4-\lambda_1\lambda_2} = \lambda_1(T_{a,1}f)(\boldsymbol{\lambda}) + \lambda_2(T_{a,2}f)(\boldsymbol{\lambda})$$

so the tuple $(T_{a,1}, T_{a,2})$ solves the Gleason problem on $\mathcal{M}$. As in the previous example, take the matrices

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad A_{a,1} = \begin{bmatrix} 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 \\ a & 0 & 0 \end{bmatrix}, \quad A_{a,2} = \begin{bmatrix} 0 & 0 & \frac{1}{4} \\ 1-a & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where $A_{a,1}$ and $A_{a,2}$ are the matrices of $T_{a,1}$ and $T_{a,2}$ with respect to the basis $\left\{\frac{4}{4-\lambda_1\lambda_2}, \frac{\lambda_1}{4-\lambda_1\lambda_2}, \frac{\lambda_2}{4-\lambda_1\lambda_2}\right\}$ of $\mathcal{M}$ and C is the matrix of the operator $G : \mathcal{M} \to \mathcal{Y}$ defined in (3.62). For every choice of a,

$$C(I - \lambda_1 A_{a,1} - \lambda_2 A_{a,2})^{-1} = \begin{bmatrix} \frac{4}{4-\lambda_1\lambda_2} & \frac{\lambda_1}{4-\lambda_1\lambda_2} & \frac{\lambda_2}{4-\lambda_1\lambda_2} \end{bmatrix},$$

which realizes $\mathcal{M}$ as the range of the observability operator of a pair $(C, \mathbf{A}_a)$. Different choices of a in (3.64) lead to nonequivalent realizations of $\mathcal{M}$. Note that $A_{a,1}$ and $A_{a,2}$ never commute which is not surprising since $\mathcal{M}$ is not backward-shift invariant as has been established in Example 3.19.

3.4 Applications of observability operators: the commutative setting

In this subsection we discuss applications of observability operators for the commutative setting. This subsection parallels Subsection 2.3.

For subspaces of $\mathcal{H}_{\mathcal{Y}}(k_d)$ invariant under the forward shift operator-tuple $M_{\boldsymbol{\lambda}}$, we have the following analogue of the Beurling–Lax–Halmos–de Branges theorem due originally to Arveson (1998) and McCullough and Trent (2000) (for the case of isometric inclusion); in fact, one can check that our proof, namely, the commutative adaptation of the proof of Theorem 2.14, follows that of Arveson (2000) if one makes the substitution $L = (\widehat{\mathcal{O}}^{\mathbf{a}}_{D_{\mathbf{T}^*}, \mathbf{T}^*})^*$ (where L is the key operator appearing in Arveson, 2000). In general, an operator Θ between two Arveson spaces $\mathcal{H}_{\mathcal{U}}(k_d)$ and $\mathcal{H}_{\mathcal{Y}}(k_d)$ is said to be *multiplier* if Θ intertwines the respective coordinate-function multipliers:

$$\theta M_{\lambda_j} f = M_{\lambda_j} \theta f \quad \text{for all} \quad f \in \mathcal{H}_{\mathcal{U}}(k_d).$$

It is straightforward to see that a multiplier Θ necessarily has the form

$$\Theta = M_\theta : f(z) \to \theta(z) \cdot f(z),$$

where $\theta(z) = \sum_{\mathbf{n}\in\mathbb{Z}^d_+} \theta_{\mathbf{n}} z^{\mathbf{n}}$ is a bounded, holomorphic $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued function on $\mathbb{B}^d$, but not all bounded, holomorphic, operator-valued functions on $\mathbb{B}^d$ are multipliers (see, e.g., Agler & McCarthy, 2002). In case the multiplication operator has operator norm at most 1, we say that θ is a *contractive multiplier* and belongs to the (commutative) multivariable Schur-class $\mathcal{S}_d(\mathcal{U}, \mathcal{Y})$. Unlike the convention in the classical case, such a multiplier θ is said to be *inner* if in addition M_θ is a partial isometry.

Theorem 3.25

(1) *A Hilbert space $\mathcal{M}$ is such that*
  (a) *$\mathcal{M}$ is contractively contained in $\mathcal{H}_\mathcal{Y}(k_d)$,*
  (b) *$\mathcal{M}$ is invariant under the Arveson-shift d-tuple $\mathbf{M}_{\boldsymbol{\lambda}}$,*
  (c) *the d-tuple*

$$\mathbf{M}_{\mathcal{M}\boldsymbol{\lambda}} = (M_{\mathcal{M},\lambda_1}, \ldots, M_{\mathcal{M},\lambda_d}),\ \textit{where}\ M_{\mathcal{M},\lambda_j} := M_{\lambda_j}|_\mathcal{M}\ \textit{for}\ j = 1, \ldots, d$$

  is a row contraction:

$$M_{\mathcal{M},\lambda_1}(M_{\mathcal{M},\lambda_1})^* + \cdots + M_{\mathcal{M},\lambda_d}(M_{\mathcal{M},\lambda_d})^* \leq I_\mathcal{M},$$

  and
  (d) *$(\mathbf{M}_{\mathcal{M}\boldsymbol{\lambda}})^*$ is strongly stable, i.e.,*

$$\sum_{\mathbf{n}\in\mathbb{Z}^d_+ :\ |\mathbf{n}|=N} \frac{|\mathbf{n}|!}{\mathbf{n}!} \left\| \mathbf{M}_{\mathcal{M},\boldsymbol{\lambda}})^{*\mathbf{n}} f \right\|^2_\mathcal{M} \to 0\ \textit{as}\ n \to \infty\ \textit{for all}\ f \in \mathcal{M}$$

  if and only if there is a coefficient Hilbert space $\mathcal{U}$ and a contractive multiplier $\theta \in \mathcal{S}_d(\mathcal{U}, \mathcal{Y})$ so that $\mathcal{M} = \theta \cdot \mathcal{H}_\mathcal{U}(k_d)$ with lifted norm

$$\|\theta \cdot f\|_\mathcal{M} = \|Qf\|_{\mathcal{H}_\mathcal{U}(k_d)}$$

  where Q is the orthogonal projection onto $(\operatorname{Ker} M_\theta)^\perp \subset \mathcal{H}_\mathcal{U}(k_d)$.

(2) *The subspace $\mathcal{M}$ in part (1) above is isometrically contained in $\mathcal{H}_\mathcal{Y}(k_d)$ if and only if the corresponding contractive multiplier $\theta \in \mathcal{S}_d(\mathcal{U}, \mathcal{Y})$ can be taken to be inner.*

Proof The proof is a straightforward commutative adaptation of the proof of Theorem 2.14 and hence will be left to the reader. We remark that, for the case where $\mathcal{M}$ is contained isometrically in $\mathcal{H}_\mathcal{Y}(k_d)$, we are unable to obtain a representer θ for which M_θ is isometric but rather only a representer with M_θ partially isometric. Indeed, one can check that the argument in the proof of Theorem 2.14 breaks down because, for the case here, M_{λ_j} is only contractive rather than isometric. □

Remark 3.26 As observed in Arveson (2002), from the function-theory point of view Theorem 3.25 is not a true analogue of the classical Beurling–Lax theorem since the characterization of θ is purely operator-theoretic with no information on the boundary behavior of the associated multiplier $\theta(z)$. This deficiency has now been remedied in the paper of Greene, Richter, and Sundberg (2002).

The following is the analogue of Theorem 2.15; we omit the proof as it exactly parallels the proof of Theorem 2.15. The result goes back to Drury (1978).

Theorem 3.27 *Suppose that $\mathbf{T} = (T_1, \ldots, T_d)$ is a commutative row-contractive operator-tuple with $\mathbf{T}^*$ strongly stable and define the defect operator $D_{\mathbf{T}^*}$ and the coefficient space $\mathcal{Y}$ as in* (2.63)*. Then there is a subspace $\mathcal{M} \subset \mathcal{H}_\mathcal{Y}(k_d)$ invariant for the backward shift operator-tuple $M^*_{\boldsymbol{\lambda}}$ on $\mathcal{H}_\mathcal{Y}(k_d)$ so that $\mathbf{T}$ is unitarily equivalent to $P_\mathcal{M} M^*_{\boldsymbol{\lambda}}|_\mathcal{M}$. In particular, $\mathbf{T}$ has an Arveson-shift dilation unitarily equivalent to $M_{\boldsymbol{\lambda}}$ on $\mathcal{H}_\mathcal{Y}(k_d)$.*

As a corollary of this result one can arrive at the von Neumann inequality

$$\|p(T_1, \ldots, T_d)\| \leq \|p(M_{\lambda_1}, \ldots, M_{\lambda_d})\|$$

of Drury (1978) and Arveson (1998) (see Remark 2.16 for the noncommutative case).

Remark 3.28 The result in Theorem 3.27 is tied to the unit ball with associated multivariable resolvent operator $(I - \lambda_1 T_1^* - \cdots - \lambda_d T_d^*)^{-1}$, associated defect operator $D_{\mathbf{T}^*} = (I - T_1 T_1^* - \cdots - T_d T_d^*)^{1/2}$, associated observability operator of the form $\widehat{\mathcal{O}}^{\mathbf{a}}_{D_{\mathbf{T}^*}, \mathbf{T}^*} = D_{\mathbf{T}^*}(I - \lambda_1 T_1^* - \cdots - \lambda_d T_d^*)^{-1}$ and associated ambient kernel function $k(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = 1/(1 - \lambda_1 \overline{\zeta_1} - \cdots - \lambda_d \overline{\zeta_d})$. We mention that there has been a lot of work centering around other types of kernels and giving a model theory for other classes of operator-tuples by using appropriately modified observability-like operators. Specifically, Müller and Vasilescu (1993) for the commutative ball case with $k(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = 1/(1 - \lambda_1 \overline{\zeta_1} - \cdots - \lambda_d \overline{\zeta_d})^m$, Curto and Vasilescu (1993, 1995) for the commutative polydisk case with $k(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = (1/(1 - \lambda_1 \overline{\zeta_1}) \cdots (1 - \lambda_d \overline{\zeta_d}))^m$, and Pott (1999) and Bhattacharyya and Sarkar (2006) for the commutative case with $k(\boldsymbol{\lambda}, \boldsymbol{\zeta}) = 1/(1 - P(\lambda_1 \overline{\zeta_1}, \ldots, \lambda_d \overline{\zeta_d}))$ with P equal to a "positively regular polynomial". The most general form of results along this line is due to Ambrozie, Engliš, and Müller (2002) and Arazy and Engliš (2003): *given a positive-definite kernel $k(\boldsymbol{\lambda}, \boldsymbol{\zeta})$ on a domain $\mathcal{D} \subset \mathbb{C}^d$ and a d-tuple of operators $\mathbf{T} = (T_1, \ldots, T_d)$ with Taylor spectrum contained in $\overline{\mathcal{D}}$ for which one can make sense of the defect operator $D_{\mathbf{T}^*} := \frac{1}{k}(T, T)$ and of the observability operator*

$$\mathcal{O}_{D_{\mathbf{T}^*}, \mathbf{T}^*} \colon x \mapsto D_{\mathbf{T}^*} k(\boldsymbol{\lambda}, T)$$

(for example, if $k(\boldsymbol{\lambda}, \boldsymbol{\zeta})$ has no zeros in $\mathcal{D} \times \mathcal{D}$ and $\mathbf{T}$ has Taylor spectrum contained in $\mathcal{D}$)*, then, under the assumption that $D_{\mathbf{T}^*} \geq 0$ and that an additional stability condition on $\mathbf{T}^*$ holds, $\mathcal{O}_{D_{\mathbf{T}^*}, \mathbf{T}^*}$ implements a unitary equivalence between $\mathbf{T}$ and $P_{\mathcal{M}} M_{\boldsymbol{\lambda}}|_{\mathcal{M}}$,* where

$$\mathcal{M} = \operatorname{Ran} \mathcal{O}_{D_{\mathbf{T}^*}, \mathbf{T}^*} \subset \mathcal{H}(k) \otimes \mathcal{Y} \text{ with } \mathcal{Y} := \overline{\operatorname{Ran}} D_{\mathbf{T}^*},$$

where $M_{\boldsymbol{\lambda}} = (M_{\lambda_1}, \ldots, M_{\lambda_d})$ is the operator-tuple of multiplication by the coordinate functions on $\mathcal{H}(k) \otimes \mathcal{Y}$, and where $\mathcal{M}$ is invariant under each of $M_{\lambda_1}^*, \ldots, M_{\lambda_d}^*$. The noncommutative case is not as well developed at this writing, but there is the paper of Popescu (1999) which handles the case of a Cartesian product of noncommutative balls (and therefore including a noncommutative polydisk). We expect that many of the ideas of the present paper, including the interplay between the noncommutative and commutative settings and the connections with system theory, have some parallels, in these other situations.

References

Agler, J., & McCarthy, J.E. (2002). *Pick interpolation and Hilbert function spaces*, Graduate Studies in Mathematics, Volume 44, Providence: American Mathematical Society.

Alpay, D., Dijksma, A., & Rovnyak, J. (2003). A theorem of Beurling-Lax type for Hilbert spaces of functions analytic in the unit ball. *Integral Equations and Operator Theory*, *47*(3), 251–274.

Alpay, D., & Dubi, C. (2005). On commuting operators solving Gleason's problem. *Proceedings of the American Mathematical Society*, *133*(11), 3285–3293.

Ambrozie, C.-G., Engliš, M., & Müller, V. (2002). Operator tuples and analytic models over general domains in $\mathbb{C}^n$. *Jounrnal of Operator Theory*, *47*, 287–302.

Arazy, J., & Engliš, M. (2003). Analytic models for commuting operator tuples on bounded symmetric domains. *Transactions of the American Mathematical Society*, *355*(2), 837–864.

Arias, A., & Popescu, G. (2000). Non-commutative interpolation and Poisson transforms. *Israel Journal of Mathematics*, *115*, 205–234.

Arveson, W. (1998). Subalgebras of C^* algebras III: Multivariable operator theory. *Acta Mathematica*, *181*, 159–228.

Arveson, W. (2000). The curvature invariant of a Hilbert module over $\mathbb{C}[z_1, \ldots, z_d]$. *Journal für die Reine und Angewandte Mathematik*, *522*, 173–236.

Ball, J.A., Bolotnikov, V., & Fang, Q. (2007a). Transfer-function realization for multipliers of the Arveson space. *Journal of Mathematical Analysis and Applications*, (to appear).

Ball, J.A., Bolotnikov, V., & Fang, Q. (2007b). Schur-class multipliers on the Fock space: de Branges-Rovnyak reproducing kernel spaces and transfer-function realizations, in *Teberiu Constantinescu Memorial Volume*, Bucharest: Theta (to appear).

Ball, J.A., Groenewald, G., & Malakorn, T. (2005). Structured noncommutative multidimensional linear system. *SIAM Journal on Control and Optimization*, *44*(4), 1474–1528.

Ball, J.A., Groenewald, G., & Malakorn, T. (2006). Conservative structured noncommutative multidimensional linear systems. In D. Alpay & I. Gohberg (Eds.), *The state space method: Generalizations and applications*. (pp. 179–223), OT 161. Basel: Birkhäuser.

Ball, J.A., & Kriete, T.L. (1987). Operator-valued Nevanlinna-Pick kernels and the functional models for contraction operators. *Integral Equations and Operator Theory, 10*(1), 17–61.

Ball, J.A., Sadosky, C., & Vinnikov, V. (2005). Conservative input-state-output systems with evolution on a multidimensional integer lattice. *Multidimensional System Signal Processes*, *16*(2), 133–198.

Ball, J.A., & Vinnikov, V. (2003). Formal reproducing kernel Hilbert spaces: The commutative and noncommutative settings. In D. Alpay (Ed.), *Reproducing Kernel spaces and applications*, (pp. 77–134), OT 143. Basel: Birkhäuser.

Ball, J.A., & Vinnikov, V. (2005). Lax-Phillips scattering and conservative linear systems: A Cuntz-algebra multidimensional setting. *Memoirs of the American Mathematical Society*, *178*(837).

Bhattacharyya, T., Eschmeier, J., & Sarkar, J. (2005). Characteristic function of a pure commuting contractive tuple. *Integral Equations and Operator Theory*, *53*(1), 23–32.

Bhattacharyya, T., & Sarkar, J. (2006). Characteristic function for polynomially contractive commuting tuples. *Journal of Mathematical Analysis and Applications*, *321*(1), 242–259.

Bolotnikov, V., & Rodman, L. (2002). Finite dimensional backward shift invariant subspaces of Arveson spaces. *Linear Algebra Applications 349*, 265–282.

Bolotnikov, V., & Rodman, L. (2004). Finite dimensional backward shift invariant subspaces of a class of reproducing kernel Hilbert spaces. *Linear Multilinear Algebra*, *52*, 321–334.

de Branges, L., & Rovnyak, J. (1966). Canonical models in quantum scattering theory. In C. H. Wilcox, (Ed.) *Perturbation theory and its applications in Quantum Mechanics* New York: Wiley. (pp. 295–392).

Chalendar, I. (2003). The operator-valued Poisson kernel and its applications. *Irish Mathematical Society Bulletin 51*, 21–44.

Curto, R.E., & Vasilescu, F.H. (1993). Standard operator models in the polydisc. *Indiana University Mathematics Journal*, *42*(3), 791–810.

Curto, R.E., & Vasilescu, F.H. (1995). Standard operator models in the polydisc, II. *Indiana Univiversity Mathematics Journal 44*(3), 727–746.

Davidson, K., & Pitts, D. (1998). Nevanlinna-Pick interpolation for non-commutative analytic Toeplitz algebras. *Integral Equations and Operator Theory*, *31*(3), 321–337.

Davidson, K.R.(2001). Free semigroup algebras: a survey. In A. A. Borichev, & N. K. Nikolski, (Eds.) *Systems, approximation, singular integral operators, and related topics* (pp. 209–240), OT 129, Basel: Birkhäuser.

Douglas, R.G. (1974). Canonical models, In *Topics in operator theory*. Mathemetical Surveys, *13*, 161–218.

Drury, S.W. (1978). A generalization of von Neumann's inequality to the complex ball. *Proceedings of the American Mathematical Society*, *68*(3), 300–304.

Dym, H. (1989). *J contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*, CBMS *71*. Providence: American Mathematical Society.

Galkowski, K. (2005). Minimal state-space realization for a class of *n*D systems, In M. A. Kaashoek, S. Seatzu, & C. van der Mee (Eds.), *Recent advances in operator theory and its applications* (pp. 179–194), OT 160, Basel: Birkhäuser.

Gleason, A.M. (1964). Finitely generated ideals in Banach algebras. *Journal of Mathematics and Mechanics*, *13*, 125–132.

Greene, D., Richter, S., & Sundberg, C. (2002). The structure of inner multipliers on spaces with complete Nevanlinna-Pick kernels. *Journal of Functional Analysis*, *194*, 311–321.
Heinz, E. (1952). Ein v. Neumannsher Satz über beschränkte Operatoren im Hilbertschen Raum, *Nachrichten der Akademie der Wissenschaften in Göttingen. II Mathematisch-Physikalische Klasse*, 5–6.
Helton, J.W. (1972/1973). The characteristic functions of operator theory and electrical network realization. *Indiana University Mathematics Journal*, *22*, 403–414.
Helton, J.W. (1974). Discrete time systems, operator models and scattering theory. *Journal of Functional Analysis*, *16* 15–38.
Henkin, G.M. (1971). The approximation of functions in pseudo-convex domains and a theorem of Z.L. Leĭbenzon. *Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques, Astronomiques et Physiques*, *19*, 37–42.
McCullough, S., & Trent, T. T. (2000). Invariance subspaces and Nevanlinna-Pick kernels, *Journal of Functional Analysis*, *178*(1), 226–249.
Müller, V., & Vasilescu, F.-H. (1993). Standard models for some commuting multioperators. *Proceedings of the American Mathematical Society*, *117*(4), 979–989.
Sz.-Nagy, B., & Foiaş, C. (1970). *Harmonic analysis of operators on Hilbert space*. Amsterdam-London: North-Holland.
Popescu, G. (1989a). Models for infinite sequences of noncommuting operators. *Acta Scientiarum Mathematicarum (Szeged)*, *53*, 355–368.
Popescu, G. (1989b). Isometric dilations for infinite sequences of noncommuting operators. *Transactions of the American Mathematical Society*, *316*, 523–536.
Popescu, G. (1989c). Characteristic functions for infinite sequences of noncommuting operators. *Journal of Operator Theory*, *22*(1), 51–71.
Popescu, G. (1989d). Multi-analytic operators and some factorization theorems. *Indiana University Mathematics Journal*, *38*(3), 693–710.
Popescu, G. (1991). von Neumann inequality for $(B(\mathcal{H})^n)_1$. *Mathematica Scandinavica*, *68*, 292-304.
Popescu, G. (1995). Multi-analytic operators on Fock spaces. *Mathematische Annalen*, *30*, 31–46.
Popescu, G. (1998). Interpolation problems in several variables. *Journal of Mathematical Analysis and Applications*, *227*(1), 227–250.
Popescu, G. (1999). Poisson transforms on some C^*-algebras generated by isometries, *Journal of Functional Analysis*, *161*(1), 27–61.
Popescu, G. (2006). Operator theory on noncommutative varieties. *Indiana University Mathematics Journal, 55*(2), 389–442.
Popescu, G. Operator theory on noncommutative varieties II. *Proceedings of the American Mathematical Society* (to appear).
Pott, S. (1999). Standard models under polynomial positivity conditions. *Journal of Operator Theory*, *41*, 365–389.

Joseph A. Ball was born in Washington, DC in 1947. He received the B.S. degree in mathematics from Georgetown University, Washington, DC, in 1969, and the M.S. and Ph.D. degrees in mathematics from the University of Virginia, Charlottesville, in 1970 and 1973, respectively. He joined the faculty at Virginia Tech, Blacksburg, in 1973, where he is currently Professor of Mathematics. Short term visits elsewhere include at the University of California at San Diego, the Weizmann Institute of Science in Israel, the University of Maryland in College Park, the Vrije Universiteit in Amsterdam, the Mathematical Sciences Research Institute in Berkeley, and the Mittag-Leffler Institute in Djursholm, Sweden. His interests include operator and function theory and its applications to control and system theory. Dr. Ball is currently an Associate Editor for the journals Integral Equations and Operator Theory, Journal of Mathematical Analysis and Applications, Proceedings of the American Mathematical Society, and Complex Analysis and Operator Theory.

Vladimir Bolotnikov was born in Kharkov, Ukraine in 1962. He received the M.S. and Ph.D. degrees in mathematics from Kharkov State University in 1984 and from Ben-Gurion University, Beer-Sheva, Israel, in 1996, respectively. After postdoctoral appointments at University of Marne la Vallee in France and at the Weizmann Institute of Science in Israel, he joined the faculty at the College of William and Mary, Williamsburg, in 1998, where he is currently Associate Professor of Mathematics. His interests include operator and function theory.

Quanlei Fang was born in China in 1979. She received the B.S. degree in Applied Mathematics in 1999 and the M.S degree in Mathematics in 2002, both from Hua Zhong university of Science and Technology, Wuhan, China. From July 2002 to July 2003, she worked as a software engineer in Beijing, China. Since August 2003, she has been a Ph.D. student in the Mathematics department at Virginia Tech, Blacksburg. Her current research interests include function theory, operator theory and the applications to system and control theory.