Compl. anal. oper. theory 99 (9999), 1–54
1661-8254/99000-0, DOI 10.1007/s11785-003-0000


**Complex Analysis and Operator Theory**

# Discrete-time dichotomous well-posed linear systems and generalized Schur-Nevanlinna-Pick interpolation

Joseph A. Ball and Michael W. Raney

**Abstract.** We introduce a class of matrix-valued functions $W$ called "$L^2$-regular". In case $W$ is $J$-inner, this class coincides with the class of "strongly regular $J$-inner" matrix functions in the sense of Arov-Dym. We show that the class of $L^2$-regular matrix functions is exactly the class of transfer functions for a discrete-time dichotomous (possibly infinite-dimensional) input-state-output linear system having some additional stability properties. When applied to $J$-inner matrix functions, we obtain a state-space realization formula for the resolvent matrix associated with a generalized Schur-Nevanlinna-Pick interpolation problem.

**Mathematics Subject Classification (2000).** 47A56; 47A57, 47A48, 93B15.

**Keywords.** shift-invariant subspace, dichotomous realization, strongly regular $J$-inner function, Grassmannian approach to interpolation.

## 1. Introduction

Over the last couple of decades, spurred on by applications in control and signal processing, there have been a number of new developments in generalizing and refining the classical theory of Nevanlinna-Pick interpolation; we mention in particular the books [19, 30, 31, 32]. This work uses the tool and language of matrix- and operator-valued functions on the unit disk or the right-half plane (see [38, 40]). We focus here on the disk setting corresponding to discrete-time (rather than continuous-time) linear systems. A cornerstone of this function theory is the Beurling-Lax-Halmos (BLH) theorem (see [26, 36, 33]) which provides a representation for a shift-invariant subspace of the Hardy space $H^2$ over the unit disk $\mathbb{D}$ of any multiplicity (or, in later versions, shift-invariant subspaces of the space $L^2$ over the unit circle $\mathbb{T}$) in terms of a matrix- or operator-valued inner function (the *representer* of the invariant subspace). The first author and Helton later obtained various generalizations of this result; we mention two particular versions: (1) where the representer is required to have $J$-unitary (rather than unitary) values on the unit circle (the BLH-theorem for the Lie group $U(m,n)$ in [20]) and (2) where the representer is only required to have invertible values on $\mathbb{T}$ but is required to represent a pair of subspaces (one forward-shift-invariant, the other backward-shift-invariant) simultaneously (the BLH theorem for the Lie group $GL(n,\mathbb{C})$ in [21]). The $U(m,n)$-BLH theorem in particular is central to the Grassmannian approach to interpolation developed also in [20] and developed further in [17, 34, 35, 23]. Formulas for the representer in terms of state-space coordinates for both the $GL(n,\mathbb{C})$ and $U(m,n)$ versions of the BLH theorem and a state-space formulation of the bitangential matrix-valued Nevanlinna-Pick interpolation problem (suggested by work of Nudelman [39]) were carried out in [1, 18, 24, 25, 19], but only for the rational case. In particular, one can characterize a shift-invariant subspace $\mathcal{M}$ (or, in the rational setting, a submodule of the module of vector rational functions over the ring of scalar rational functions with no poles in a prescribed subset $\sigma$ of the extended complex plane $\mathbb{C}\cup\{\infty\}$) via pole and zero directional data which can be encoded in a five-tuple of finite matrices (an *admissible Sylvester data set* $\mathfrak{S}$); one can then build the BLH-representer for $\mathcal{M}$ in realization form via various matrix constructions from the Sylvester data set $\mathfrak{S}$.

The purpose of the present paper is to extend and complete the state-space formulas from [18, 24, 19] to handle the general nonrational setting in [20, 21]. We make precise the class of realizations (*strongly bi-dichotomous realizations*) required for state-space implementation of the $GL(n,\mathbb{C})$-BLH theorem from [21]. As was observed in [21], one gets the $U(m,n)$-BLH theorem by specializing the $GL(n,\mathbb{C})$-BLH theorem for the pair of subspaces to the particular case where $\mathcal{M}^\times = \mathcal{M}^{\perp J}$. We make use of this idea here to obtain the infinite-dimensional state-space implementation of the $U(m,n)$-BLH theorem as the special case of the $GL(n,\mathbb{C})$ case where $\mathcal{M}^\times$ is the $J$-orthogonal complement $\mathcal{M}^{\perp J}$ of $\mathcal{M}$. The results lead to new types of realization formulas which are closely related to recent work of Arova [14, 15, 16]. We also are able to identify the class of $L^2$-regular $J$-inner functions arising here as being the same as the class of *strongly Arov-regular* $J$-inner functions which has played a prominent role in the study of inverse problems for canonical systems by Arov and Dym [4, 5, 6, 7, 8, 9, 10, 11, 12].

The paper is organized as follows. In Section 2 we organize preliminary material concerning dichotomous linear systems and formally define the class of strongly bi-dichotomous realizations which plays a key role in the analysis to come. In Section 3 we introduce and carefully prove the $GL(n,\mathbb{C})$ BLH theorem from [21] to allow for even $n=\infty$ (i.e., for an infinite-dimensional coefficient space $\mathcal{U}$). In Section 4 we introduce the appropriate notion of admissible Sylvester data set for the infinite-dimensional setting, and obtain the connection between Sylvester data sets and (forward and backward) shift-invariant subspaces of $L^2\otimes\mathcal{U}$; this extends the work of [19, Part III] to a nonrational context and establishes a connection with functions having meromorphic pseudocontinuation of bounded type (see [29, 28]).

Section 5 gives the algorithm for the state-space implementation of the $GL(n,\mathbb{C})$-BLH theorem from [21]; this is a canonical generalization of the main result of [18] (see [19, Theorem 5.5.2]). In Section 6 we specialize the main result of Section 5 to the case where $\mathcal{M}^{\times} = \mathcal{M}^{\perp J}$ to obtain the state-space implementation of the $U(m,n)$-BLH theorem obtained in [20]. Section 7 draws on results from [22] to study the special structure arising when the $J$-unitary function is actually $J$-inner, i.e., the associated kernel $K_{\Theta}(z,w) = [J - \Theta(z) J \Theta(w)^*]/(1 - z\overline{w})$ is a positive kernel on $\mathbb{D} \times \mathbb{D}$. The final Section 8 makes the connection with generalized Nevanlinna-Pick interpolation and with the class of strongly Arov-regular $J$-inner functions introduced in [4].

An interesting project is to develop the analogues of all the results of the present paper for the case where the unit disk $\mathbb{D}$ is replaced by the right half plane. The difficulty for realization theory is that the point at infinity is now on the boundary of the domain rather than in the interior, but a formalism has now been worked out to handle these difficulties (see the recent book [42]). Nevertheless, the realization theory for a Schur-class function on the right half plane can be tricky and counter-intuitive—see [37].

Finally it our pleasure to thank our colleague Vladimir Bolotnikov of the College of William and Mary and the anonymous reviewer for valuable comments on the original version of this paper.

## 2. Preliminaries

### 2.1. Notational conventions

Throughout the paper we shall use the following notation, most of which is standard in the literature. We let $\mathbb{Z}$ denote the set of all integers, $\mathbb{Z}_+$ the set of nonnnegative integers and $\mathbb{Z}_- = \mathbb{Z} \setminus \mathbb{Z}_+$ the set of negative integers. For $\mathcal{U}$ a Hilbert space, we use $\ell_{\mathcal{U}}(\mathbb{Z})$, $\ell_{\mathcal{U}}(\mathbb{Z}_+)$ and $\ell_{\mathcal{U}}(\mathbb{Z}_-)$ to denote the space of norm-square-summable sequences with values in $\mathcal{U}$ indexed by $\mathbb{Z}$, $\mathbb{Z}_+$ and $\mathbb{Z}_-$ respectively with the usual Hilbert space norm

$$\|\{u(n)\}_{n\in\mathbb{Z}}\|_{\ell^2_{\mathcal{U}}(\mathbb{Z})} = \sum_{n\in\mathbb{Z}} \|u(n)\|^2_{\mathcal{U}}.$$

We view $\ell^2_{\mathcal{U}}(\mathbb{Z}_+)$ and $\ell^2_{\mathcal{U}}(\mathbb{Z}_-)$ as subspaces of $\ell^2_{\mathcal{U}}(\mathbb{Z})$ in the usual way. The $Z$-transform

$$\{u(n)\}_{n\in\mathbb{Z}} \mapsto \widehat{u}(z) := \sum_{n\in\mathbb{Z}} u(n) z^n \tag{2.1}$$

transforms the sequence spaces $\ell^2_{\mathcal{U}}(\mathbb{Z})$, $\ell^2_{\mathcal{U}}(\mathbb{Z}_+)$ and $\ell^2_{\mathcal{U}}(\mathbb{Z}_-)$ into the respective Hilbert spaces of measurable $\mathcal{U}$-valued functions on the unit circle $\mathbb{T}$

$$L^2_{\mathcal{U}} = \left\{ f(z) = \sum_{n\in\mathbb{Z}} f_n z^n \colon f_n \in \mathcal{U} \text{ with } \|f\|^2 = \sum_{n\in\mathbb{Z}} \|f_n\|^2_{\mathcal{U}} < \infty \right\},$$

$$H^2_{\mathcal{U}} = \left\{ f(z) = \sum_{n \in \mathbb{Z}_+} f_n z^n \colon f_n \in \mathcal{U} \text{ with } \|f\|^2 = \sum_{n \in \mathbb{Z}} \|f_n\|^2_{\mathcal{U}} < \infty \right\},$$

and

$$H^{2\perp}_{\mathcal{U}} = \left\{ f(z) = \sum_{n \in \mathbb{Z}_+} f_n z^n \colon f_n \in \mathcal{U} \text{ with } \|f\|^2 = \sum_{n \in \mathbb{Z}_-} \|f_n\|^2_{\mathcal{U}} < \infty \right\}.$$

When convenient, we view an element $f$ of $H^2_{\mathcal{U}}$ as a holomorphic function on the unit disk $\mathbb{D}$ and an element $g$ of $H^{2\perp}_{\mathcal{U}}$ as also an analytic $\mathcal{U}$-valued function on the exterior $\mathbb{D}_e$ of the closed unit disk (with value 0 at infinity).

The Banach space of essentially bounded measurable $\mathcal{U}$-valued functions on the unit circle is denoted by $L^\infty_{\mathcal{U}}$. We shall have use of two of its subspaces

$$H^\infty_{\mathcal{U}} = L^\infty_{\mathcal{U}} \cap H^2_{\mathcal{U}},$$
$$\overline{H^\infty_{\mathcal{U},0}} = L^\infty_{\mathcal{U}} \cap H^{2\perp}_{\mathcal{U}}$$

which can also be viewed as bounded $\mathcal{U}$-valued analytic functions on the unit disk $\mathbb{D}$ and as bounded $\mathcal{U}$-valued analytic functions on the exterior of the unit disk $\mathbb{D}_e$ with value 0 at infinity, respectively; as is suggested by the notation, as a function on the unit circle $\mathbb{T}$, $\overline{H^\infty_{\mathcal{U},0}}$ is just the image of the space $H^\infty_{\mathcal{U},0}$ of bounded $\mathcal{U}$-valued analytic functions on $\mathbb{D}$ vanishing at the origin under complex conjugation. If $\mathcal{U}$ and $\mathcal{Y}$ are two Hilbert spaces, $\mathcal{L}(\mathcal{U}, \mathcal{Y})$ denotes the space of bounded, linear operators from $\mathcal{U}$ to $\mathcal{Y}$; we shall abbreviate $\mathcal{L}(\mathcal{U},\mathcal{U})$ to $\mathcal{L}(\mathcal{U})$.

We shall use one nonstandard notation: $L^2_{\mathcal{L}(\mathcal{U},\mathcal{Y})}$ denotes the space of weakly measurable $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued functions $W(z)$ on $\mathbb{T}$ with the property that the multiplication operator$M_W \colon u \mapsto W(z) \cdot u$ maps constant vectors $u \in \mathcal{U}$ into $L^2_{\mathcal{Y}}$:

$$L^2_{\mathcal{L}(\mathcal{U},\mathcal{Y})} = \{W \colon \mathbb{T} \to \mathcal{L}(\mathcal{U}, \mathcal{Y}) \colon M_W \in \mathcal{L}(\mathcal{U}, L^2_{\mathcal{Y}})\}.$$

Note that a consequence of the closed graph theorem (see e.g. [41, statement 2.15]) is that $M_W$ is bounded as an operator from $\mathcal{U}$ into $L^2_{\mathcal{Y}}$ just from the assumption that $M_W$ maps any vector $u \in \mathcal{U}$ into $L^2_{\mathcal{Y}}$. We shall also have occasion to need a weak $L^1$-version

$$L^1_{\mathcal{L}(\mathcal{U},\mathcal{Y})} = \{W \colon \mathbb{T} \to \mathcal{L}(\mathcal{U}, \mathcal{Y}) \colon \langle Wu, y \rangle_{\mathcal{Y}} \in L^1 \text{ for } u \in \mathcal{U} \text{ and } y \in \mathcal{Y}\}.$$

Here $L^1$ is the standard Lebesgue space of modulus-integrable measurable complex-valued functions on $\mathbb{T}$. Note that in case $\mathcal{U}$ and $\mathcal{Y}$ are finite-dimensional and $\mathcal{L}(\mathcal{U}, \mathcal{Y})$ is identified with the collection of finite matrices of some fixed size, then the space $L^2_{\mathcal{L}(\mathcal{U},\mathcal{Y})}$ is simply the space of $\mathcal{L}(\mathcal{U}, \mathcal{Y})$-valued functions with each matrix entry in $L^2$; a similar remark applies to the space $L^1_{\mathcal{L}(\mathcal{U},\mathcal{Y})}$.

The following fact will be useful in the sequel.

**Proposition 2.1.** *If* $W \in L^2_{\mathcal{L}(\mathcal{U},\mathcal{Y})}$ *and* $\widetilde{W} \in L^2_{\mathcal{L}(\mathcal{X},\mathcal{Y})}$, *then* $\widetilde{W}^* W \in L^1_{\mathcal{L}(\mathcal{U},\mathcal{X})}$.

*Proof.* Suppose that $W$ and $\widetilde{W}$ are as in the hypotheses. We wish to show that $\langle \widetilde{W}^* W u, x\rangle_{\mathcal{X}} \in L^1$ for each $u \in \mathcal{U}$ and $x \in \mathcal{X}$. By duality it suffices to show that the scalar-valued function

$$z \mapsto \langle \widetilde{W}(z)^* W(z) u, f(z)x\rangle_{\mathcal{X}}$$

is in $L^1$ for each $f \in L^\infty$ and $x \in \mathcal{X}$. For this purpose, note that

$$\langle \widetilde{W}(z)^* W(z) u, f(z)x\rangle_{\mathcal{X}} = \langle W(z)u, \widetilde{W}(z) f(z) x\rangle_{\mathcal{Y}} \tag{2.2}$$

where by assumption both $W(z)u$ and $\widetilde{W}(z)f(z)x$ are in $L^2_{\mathcal{Y}}$. By the Cauchy-Schwarz inequality, the function in (2.2) is in $L^1$ as needed. □

### 2.2. Preliminaries from linear system theory

By an *output pair* we mean any pair of operators $(C, A)$ with $A \in \mathcal{L}(\mathcal{X})$ and $C \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ for a pair of Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$. Associated with any output pair $(C, A)$ is the *forward-time state-output linear system*

$$\Sigma^f_{\text{state/out}} : \left\{ \begin{array}{rcl} x(n+1) & = & Ax(n) \\ y(n) & = & Cx(n). \end{array} \right. \tag{2.3}$$

Note that the initial condition $x(0) = x_0$ uniquely determines the solution of the system equations $(x(n), y(n))$ for $n \geq 0$ according to

$$x(n) = A^n x(0), \qquad y(n) = CA^n x(0).$$

We denote by $\mathcal{O}^f_{C,A}$ the *observation operator* which assigns to $x(0)$ the resulting output sequence $\{y(n)\}_{n\geq 0}$:

$$\mathcal{O}^f_{C,A} \colon x \mapsto \{CA^n x\}_{n\in\mathbb{Z}_+}. \tag{2.4}$$

If this operator is injective, we say that the pair $(C, A)$ is *observable*. In case $\mathcal{O}^f_{C,A}$ maps $\mathcal{X}$ into $\ell^2_{\mathcal{Y}}(\mathbb{Z}_+)$, we say that the pair $(C, A)$ is *output-stable*; in this case, by the closed-graph theorem it is automatic that $\mathcal{O}^f_{C,A}$ is bounded as an operator from $\mathcal{X}$ into $\ell^2_{\mathcal{Y}}(\mathbb{Z}_+)$. If it happens that $\mathcal{O}^f_{C,A}$ is bounded below as an operator from $\mathcal{X}$ into $\ell^2_{\mathcal{Y}}(\mathbb{Z}_+)$, namely

$$\|\mathcal{O}^f_{Z,B} x\|_{\ell^2_{\mathcal{Y}}(\mathbb{Z}_+)} \geq \delta \|x\|_{\mathcal{X}} \text{ for some } \delta > 0 \text{ for all } x \in \mathcal{X},$$

we say that the output pair $(C, A)$ is *exactly observable*. In this case the *observability gramian* $\mathcal{G}_{C,A}$ given by

$$\mathcal{G}_{C,A} = (\mathcal{O}^f_{C,A})^* \mathcal{O}^f_{C,A} = \text{s-lim}_{N\to\infty} \sum_{n=0}^{N} A^{*n} C^* C A^n$$

is the unique solution $H$ of the Stein equation

$$H - A^* H A = C^* C. \tag{2.5}$$

Moreover, iteration of (2.5) gives

$$H = \sum_{n=0}^{N-1} A^{*n} C^* C A^n + A^{*N} H A^N \tag{2.6}$$

for every $N = 1, 2, \dots$. For the case where $H = \mathcal{G}_{C,A}$, the first term on the right hand side converges strongly to $H = \mathcal{G}_{C,A}$ as $N \to \infty$ and we conclude that

$$\lim_{N\to\infty} \langle \mathcal{G}_{C,A} A^N x, A^N x\rangle_{\mathcal{X}} = 0 \text{ for all } x \in \mathcal{X}.$$

As $\mathcal{G}_{C,A}$ is strictly positive definite in the exactly observable case, we get from this that

$$\lim_{N\to\infty} \|A^N x\|^2_{\mathcal{X}} = 0 \text{ for all } x, \in \mathcal{X} \tag{2.7}$$

i.e., $A$ ***is strongly stable*** whenever there is a $C$ so that the output pair $(C, A)$ is exactly observable.

Application of the $Z$-transform (2.1) leads to the $Z$-transformed version $\widehat{\mathcal{O}}^f_{C,A}$ of the observation operator $\mathcal{O}^f_{C,A}$

$$\widehat{\mathcal{O}}^f_{C,A} \colon x \mapsto \widehat{\mathcal{O}^f_{C,A} x} = \sum_{n=0}^{\infty} (CA^n x) z^n = C(I - zA)^{-1} x,$$

where the latter equality holds at least in a neighborhood of 0. In particular, if $(C, A)$ is output-stable, then $\widehat{\mathcal{O}}^f_{C,A}$ maps $\mathcal{X}$ into $H^2_{\mathcal{Y}}$.

Given the output pair $(C, A)$, instead of the forward-time state-output linear system (2.3) we could have considered the *backward-time state-output linear system*

$$\Sigma^b_{\text{state/out}} \colon \left\{ \begin{array}{lcl} x(n) & = & Ax(n+1) \\ y(n) & = & Cx(n+1). \end{array} \right. \tag{2.8}$$

In this case, specifying an initial condition $x(0) = x_0$ and letting the system run determines $(x(n), y(n))$ for $n \in \mathbb{Z}_- = \{n \in \mathbb{Z} \colon n < 0\}$ according to

$$x(-n) = A^n x(0), \qquad y(-n) = CA^{n-1} x(0) \text{ for } n = 1, 2, \dots.$$

We therefore define the associated *backward-time observation operator* by

$$\mathcal{O}^b_{C,A} \colon x \mapsto \{CA^{-n} x\}_{n \in \mathbb{Z}_-}$$

with corresponding $Z$-transformed version

$$\widehat{\mathcal{O}}^b_{C,A} \colon x \mapsto \sum_{n=1}^{\infty} (CA^{n-1} x) z^{-n} = C(zI - A)^{-1} x$$

where the latter equality holds at least in a neighborhood of infinity. Note that the pair is *output-stable* for the backward-time system (2.8) in the sense that $\mathcal{O}^b_{C,A}$ maps $\mathcal{X}$ into $\ell^2_{\mathcal{Y}}(\mathbb{Z}_-)$ if and only if $(C, A)$ is output-stable for the forward-time system in the sense defined above, namely,

$$\sum_{n=0}^{\infty} \|CA^n x\|^2_{\mathcal{Y}} < \infty \text{ for all } x \in \mathcal{X}. \tag{2.9}$$

Thus output-stability for the pair $(C, A)$ is also equivalent to the the $Z$-transformed backward-time observation operator $\widehat{\mathcal{O}}^b_{C,A}$ mapping $\mathcal{X}$ into $H^{2\perp}_{\mathcal{Y}}$.

All these ideas have dual formulations on the input side. By an *input pair* we mean any pair of operators $(Z, B)$ where $Z \in \mathcal{L}(\mathcal{X})$ and $B \in \mathcal{L}(\mathcal{U}, \mathcal{X})$ for some Hilbert spaces $\mathcal{X}$ and $\mathcal{U}$. Associated with any input pair $(Z, B)$ is a *forward-time input-state linear system* $\Sigma^f_{\text{in/state}}$ given by

$$x(n+1) = Zx(n) + Bu(n). \tag{2.10}$$

If $\{u(n)\}_{n\in\mathbb{Z}_-}$ is a $\mathcal{U}$-valued input string finitely supported on $\mathbb{Z}_-$ (say $u(n) = 0$ for $n < -N$ for some $N < \infty$) and if we assume that $x(n) = 0$ for all $n \le -N$, then we may use the system equation (2.10) to solve for $x(0)$:

$$x(0) = \sum_{n=0}^{N} Z^{n-1}Bu(-n) = \begin{bmatrix} Z^{N-1}B & Z^{N-2}B & \cdots & ZB & B \end{bmatrix} \begin{bmatrix} u(-N) \\ u(-N+1) \\ \vdots \\ u(-2) \\ u(-1) \end{bmatrix}.$$

We therefore define the *forward-time control operator* $\mathcal{C}^f_{Z,B}$ be the block row-matrix

$$\mathcal{C}^f_{Z,B} = \text{row}_{n\in\mathbb{Z}_-}[Z^{-n-1}B]$$

considered as an operator from the space of finitely-supported $\mathcal{U}$-valued sequences on $\mathbb{Z}_-$ into $\mathcal{X}$. If this operator has dense range in $\mathcal{X}$, we say that the pair $(Z, B)$ is *controllable*. In case $\mathcal{C}^f_{Z,B}$ extends to define a bounded operator from $\ell^2_{\mathcal{Y}}(\mathbb{Z}_-)$ into $\mathcal{X}$, we say that the input-pair $(Z, B)$ is *input-stable*. In case this extended operator maps $\ell^2_{\mathcal{Y}}$ onto the state space $\mathcal{X}$, we say that $(Z, B)$ is *exactly controllable*. In the exactly controllable case, the *controllability gramian*

$$\mathcal{G}_{Z,B} := \mathcal{C}^f_{Z,B}(\mathcal{C}^f_{Z,B})^* = \text{s-lim}_{N\to\infty} \sum_{n=0}^{N} Z^n BB^* Z^{*n}$$

is strictly positive-definite on $\mathcal{X}$ and is the unique solution $H$ of the control Stein equation

$$H - ZHZ^* = BB^*. \tag{2.11}$$

Iteration of (2.11) gives

$$H = \sum_{n=0}^{N-1} Z^n BB^* Z^{*n} + Z^N HZ^{*N}. \tag{2.12}$$

From this we are able to conclude that $Z^*$ is *strongly stable*, i.e.,

$$\lim_{N\to\infty} \|Z^{*N}x\|^2_{\mathcal{X}} = 0 \text{ for all } x \in \mathcal{X} \tag{2.13}$$

whenever there is an input operator $B$ so that the input pair $(Z, B)$ is exactly controllable.

We define the $Z$-transformed forward-time control operator $\widehat{\mathcal{C}}^f_{Z,B}$ to map polynomials in $z^{-1}$ with coefficients in $\mathcal{U}$ into $\mathcal{X}$ by

$$\widehat{\mathcal{C}}^f_{Z,B} \colon \sum_{n=1}^{N} u(n) z^{-n} \mapsto \sum_{n=1}^{N} Z^{n-1} B u(n).$$

In the input-stable case, $\widehat{\mathcal{C}}^f_{Z,B}$ extends to define a bounded operator from $H^{2\perp}_{\mathcal{U}}$ into $\mathcal{X}$ according to the formula

$$\widehat{\mathcal{C}}^f_{Z,B} \colon \sum_{n=1}^{\infty} u(n) z^{-n} \mapsto \sum_{n=1}^{\infty} Z^{n-1} B u(n).$$

This extended operator has dense range in $\mathcal{X}$ (respectively, has range equal to all of $\mathcal{X}$) if the pair $(Z, B)$ is controllable (respectively, is exactly controllable).

Given an input pair $(Z, B)$ rather than the forward-time input-state system (2.10) we could have considered instead the *backward-time input-state system* $\Sigma^b_{\text{in/state}}$ given by

$$x(n) = Z x(n+1) + B u(n). \tag{2.14}$$

Given a $\mathcal{U}$-valued input string $\{u(n)\}_{n \in \mathbb{Z}_+}$ finitely supported on $\mathbb{Z}_+$ (say $u(n) = 0$ for all $n > N$ for some $N < \infty$), under the assumption that the state $x(n)$ also vanishes for $n > N$, we can compute $x(0)$ by iterating the system equation (2.14) according to

$$x(0) = \sum_{n=0}^{N} Z^n B u(n) = \begin{bmatrix} B & ZB & \cdots & Z^{N-1}B & Z^N B \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N-1) \\ u(N) \end{bmatrix}.$$

This suggests that we define the *backward-time control operator* $\mathcal{C}^b_{Z,B}$ as the block row-matrix

$$\mathcal{C}^b_{Z,B} = \text{row}_{n \in \mathbb{Z}_+} [Z^n B]$$

considered as an operator mapping $\mathcal{U}$-valued input strings finitely supported on $\mathbb{Z}_+$ into $\mathcal{X}$. In case the pair $(Z, B)$ is input-stable, then $\mathcal{C}^f_{Z,B}$ extends to define a bounded operator from $\ell^2_{\mathcal{U}}(\mathbb{Z}_+)$ into $\mathcal{X}$. The $Z$-transformed version $\widehat{\mathcal{C}}^b_{Z,B}$ is defined to take polynomials with coefficients in $\mathcal{U}$ into $\mathcal{X}$ according to the formula

$$\widehat{\mathcal{C}}^b_{Z,B} \colon \sum_{n=0}^{\infty} u(n) z^{n} \mapsto \sum_{n=0}^{\infty} Z^{n} B u(n).$$

In the input-stable case, $\widehat{\mathcal{C}}^b_{Z,B}$ extends to define a bounded linear operator from $H^2_{\mathcal{U}}$ into $\mathcal{X}$.

Given any output pair $(C, A)$ we may consider $(Z, B) := (A^*, C^*)$ as an input pair, and given an input pair $(Z, B)$ we may consider $(C, A) = (B^*, Z^*)$ as an output pair. Note that each of the concepts discussed for an output pair transfers to the corresponding concept for an input pair under this duality. Thus, an output pair $(C, A)$ is output-stable if and only if $(A^*, C^*)$ is input-stable, the input pair $(Z, B)$ is input-stable if and only if the output pair $(B^*, Z^*)$ is output stable, and then

$$(\widehat{\mathcal{O}}^f_{C,A})^* = \widehat{\mathcal{C}}^b_{A^*,C^*} \in \mathcal{L}(\mathcal{X}, H^2_{\mathcal{Y}}), \qquad (\widehat{\mathcal{O}}^b_{C,A})^* = \widehat{\mathcal{C}}^f_{A^*,C^*} \in \mathcal{L}(\mathcal{X}, H^{2\perp}_{\mathcal{Y}}),$$
$$(\widehat{\mathcal{C}}^f_{Z,B})^* = \widehat{\mathcal{O}}^b_{B^*,Z^*} \in \mathcal{L}(\mathcal{X}, H^{2\perp}_{\mathcal{U}}), \qquad (\widehat{\mathcal{C}}^b_{Z,B})^* = \widehat{\mathcal{O}}^f_{B^*,Z^*} \in \mathcal{L}(\mathcal{X}, H^2_{\mathcal{U}}).$$

Moreover, the output pair $(C, A)$ is observable (respectively, exactly observable) if and only if the input pair $(A^*, C^*)$ is controllable (respectively, exactly controllable).

We now introduce the notion of dichotomous discrete-time linear system. A *descriptor system* (in output-nulling form—see [13]) is a system of the form

$$\begin{aligned} Ex(n+1) &= Ax(n) + B_i u(n) + B_o y(n) \\ 0 &= Cx(n) + D_i u(n) + D_o y(n) \end{aligned} \tag{2.15}$$

The associated *extended behavior* $\mathcal{W}_e$ is the set of all triples of functions

$$\{u(n), x(n), y(n)\}_{n\in\mathbb{Z}}$$

(with $u(n) \in \mathcal{U}$, $x(n) \in \mathcal{X}$, $y(n) \in \mathcal{Y}$ for all $n \in \mathbb{Z}$) such that the system of equations (2.15) is satisfied. The *behavior* $\mathcal{W}$ of the system (2.15) consists of all input-output pairs $\{u(n), y(n)\}_{n\in\mathbb{Z}}$ for which there exists a state trajectory $\{x(n)\}_{n\in\mathbb{Z}}$ so that $\{u(n), x(n), y(n)\}_{n\in\mathbb{Z}}$ is an element of the extended behavior $\mathcal{W}_e$. We say that the system (2.15) has *dichotomy* (with respect to the pair $\mathcal{T}/L^2$) if the following property holds: *for each choice of finitely supported input signal $\{u(n)\}_{n\in\mathbb{Z}}$ (so $\widehat{u}(z) = \sum_{n\in\mathbb{Z}} u(n)z^n$ is in the space $\mathcal{T}_{\mathcal{U}}$ of trigonometric polynomials with coefficients in $\mathcal{U}$), there is a unique choice of output signal $\{y(n)\}_{n\in\mathbb{Z}}$ so that $\{y(n)\}_{n\in\mathbb{Z}} \in \ell^2_{\mathcal{Y}}(\mathbb{Z})$ (and hence $\widehat{y}(z) = \sum_{n\in\mathbb{Z}} y(n)z^n \in L^2_{\mathcal{Y}}$) and* $\{u(n), y(n)\}_{n\in\mathbb{Z}} \in \mathcal{W}$. In case the corresponding state trajectory $\{x(n)\}_{n\in\mathbb{Z}}$ is also uniquely determined, one would say that the system (2.15) is an *observable dichotomous system*. One way for this to happen is that there exists a decomposition of the state space $\mathcal{X} = \mathcal{X}_+ \dot{+} \mathcal{X}_-$ so that the operators in (2.15) have matrix decompositions of the form

$$E = \begin{bmatrix} I_{\mathcal{X}_+} & 0 \\ 0 & -A_- \end{bmatrix}, \qquad A = \begin{bmatrix} A_+ & 0 \\ 0 & -I_{\mathcal{X}_-} \end{bmatrix}, \qquad B_i = \begin{bmatrix} B_+ \\ B_- \end{bmatrix}, \qquad B_o = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$
$$C = \begin{bmatrix} C_+ & C_- \end{bmatrix}, \qquad D_i = D, \qquad D_o = -I_{\mathcal{Y}}$$

from which it follows that the system equations (2.15) assume the form

$$\begin{aligned} x_+(n+1) &= A_+ x_+(n) + B_+ u(n) \\ x_-(n) &= A_- x_-(n+1) + B_- u(n) \\ y(n) &= C_+ x_+(n) + C_- x_-(n) + Du(n) \end{aligned} \tag{2.16}$$

where

1. either $(C_+, A_+)$ is an output-stable pair or $(A_+, B_+)$ is an input-stable pair (so $W_+(z) = C_+(I_{\mathcal{X}_+} - zA_+)^{-1}B_+ \in H^2_{\mathcal{L}(\mathcal{U},\mathcal{Y})}$), and
2. either $(C_-, A_-)$ is an output-stable pair or $(A_-, B_-)$ is an input stable pair (so $W_-(z) = C_-(zI_{\mathcal{X}_-} - A_-)^{-1}B_- \in H^{2\perp}_{\mathcal{L}(\mathcal{U},\mathcal{Y})}$).

(see Section 2.1 for our notational conventions). Indeed, in this situation, given a input signal $\{u(n)\}_{n\in\mathbb{Z}}$ of finite support, say on $-N \le n \le N$, and an initialization of the state by $x_+(-N) = 0$ and $x_-(N+1) = 0$, one can use the first of equations (2.16) to recursively compute $x_+(n)$ for $n > -N$ and the second of equations (2.16) to compute $x_-(n)$ for $n \le N$. If we then set $x_+(n) = 0$ for $n < -N$ and $x_-(n) = 0$ for $n > N+1$, then the first two of equations (2.16) are satisfied for all $n \in \mathbb{Z}$. If we then use the last of equations (2.16) to compute $\{y(n)\}_{n\in\mathbb{Z}}$, apply the formal $Z$-transform $\{x(n)\}_{n\in\mathbb{Z}} \mapsto \sum_{n\in\mathbb{Z}} x(n)z^n$ in all three of equations (2.16), and then eliminate the state variable, we arrive at $\widehat{y}(z) = W(z)\widehat{u}(z)$ where the *transfer function* $W(z)$ is given by

$$W(z) = zC_+(I - zA_+)^{-1}B_+ + D + C_-(zI - A_-)^{-1}B_-.$$

If both conditions (1) and (2) above hold, then $W(z) = W_+(z) + D + W_-(z)$ is in $L^2_{\mathcal{L}(\mathcal{U},\mathcal{Y})}$ and it follow that $M_W$ maps trigonometric polynomials $\mathcal{T}_\mathcal{U}$ into $L^2_\mathcal{Y}$.

The particular form of dichotomy which we shall work with here is as follows. We shall say that the system is *strongly dichotomous* if both $(C_+, A_+)$ and $(C_-, A_-)$ are exactly observable, $(A_+, B_+)$ and $(A_-, B_-)$ are both controllable, and both $A_+$ and $A_-$ are strongly bi-stable.

Analogously, we say that system (2.15) has *strong inverse-dichotomy* (with respect to the pair $\mathcal{T}/L^2$) if the following property holds: *for each choice of finitely supported output trajectory $\{y(n)\}_{n\in\mathbb{Z}}$ (so $\widehat{y}(z) = \sum_{n\in\mathbb{Z}} y(n)z^n$ is in the trigonometric-polynomial space $\mathcal{T}_\mathcal{Y}$), there is a unique choice of input trajectory $\{u(n)\}_{n\in\mathbb{Z}}$ so that $\{u(n)\}_{n\in\mathbb{Z}} \in \ell^2_\mathcal{U}(\mathbb{Z})$ (and hence $\widehat{u}(z) = \sum_{n\in\mathbb{Z}} u(n)z^n \in L^2_\mathcal{U}$) and additionally* $\{u(n), y(n)\}_{n\in\mathbb{Z}} \in \mathcal{W}$. One way for this to happen is that there is a decomposition $\mathcal{X}^\times = \mathcal{X}^\times_+ \dot{+} \mathcal{X}^\times_-$ so that the operators in (2.15) take on the form

$$E = \begin{bmatrix} I_{\mathcal{X}^\times_+} & 0 \\ 0 & -A^\times_- \end{bmatrix}, \quad A = \begin{bmatrix} A^\times_+ & 0 \\ 0 & -I_{\mathcal{X}^\times_-} \end{bmatrix}, \quad B_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad B_o = \begin{bmatrix} B^\times_+ \\ B^\times_- \end{bmatrix}$$
$$C = \begin{bmatrix} C^\times_+ & C^\times_- \end{bmatrix}, \quad D_i = -I_\mathcal{U}, \quad D_o = D^\times$$

so the system equations have the form

$$\begin{aligned} x^\times_+(n+1) &= A^\times_+ x^\times_+(n) + B^\times_+ y(n) \\ x^\times_-(n) &= A^\times_- x^\times_-(n+1) + B^\times_- y(n) \\ u(n) &= C^\times_+ x^\times_+(n) + C^\times_- x^\times_-(n) + D^\times y(n) \end{aligned} \tag{2.17}$$

where

1. either $(C^\times_+, A^\times_+)$ is an output-stable pair or $(A^\times_+, B^\times_+)$ is an input-stable pair (so $W^\times_+(z) = C^\times_+(I_{\mathcal{X}^\times_+} - zA^\times_+)^{-1}B^\times_+ \in H^2_{\mathcal{L}(\mathcal{Y},\mathcal{U})}$), and

2. either $(C_-^\times, A_-^\times)$ is an output-stable pair or $(A_-^\times, B_-^\times)$ is an input stable pair (so $W_-^\times(z) = C_-^\times(zI_{\mathcal{X}_-^\times} - A_-^\times)^{-1}B_-^\times \in H^{2\perp}_{\mathcal{L}(\mathcal{Y},\mathcal{U})}$).

The particular form of dichotomy which we shall impose for the inverse system is as follows. We shall say that the system is *strongly inverse-dichotomous* if both $(A_+^\times, B_+^\times)$ and $(A_-^\times, B_-^\times)$ are exactly controllable, both $(C_+^\times, A_+^\times)$ and $C_-^\times, A_-^\times)$ are observable, and both $A_+^\times$ and $A_-^\times$ are strongly bi-stable.

If the system is both strongly dichotomous and strongly inverse-dichotomous, we shall say that the system is *strongly bi-dichotomous*. For convenience in the sequel we shall take $\mathcal{U} = \mathcal{Y}$ without loss of generality. We now summarize in a formal definition what we mean by a strongly bi-dichotomous realization for a given $\mathcal{L}(\mathcal{U})$-valued function $W$ on $\mathbb{T}$.

**Definition 2.2.** *Let $W$ be a given $\mathcal{L}(\mathcal{U})$-valued function on $\mathbb{T}$. We say that $W$ has a* strongly bi-dichotomous realization *if there exist operators*

$$C_\pm \in \mathcal{L}(\mathcal{X}_\pm, \mathcal{U}),\ A_\pm \in \mathcal{L}(\mathcal{X}_\pm),\ B_\pm \in \mathcal{L}(\mathcal{U}, \mathcal{X}_\pm),\ D \in \mathcal{L}(\mathcal{U}),$$
$$C_\pm^\times \in \mathcal{L}(\mathcal{X}_\pm^\times, \mathcal{U}),\ A_\pm^\times \in \mathcal{L}(\mathcal{X}_\pm^\times),\ B_\pm^\times \in \mathcal{L}(\mathcal{U}, \mathcal{X}_\pm^\times), D^\times \in \mathcal{L}(\mathcal{U})$$

*subject to*

1. *$(C_+, A_+)$ and $(C_-, A_-)$ are exactly observable,*
2. *$(A_+, B_+)$ and $(A_-, B_-)$ are controllable,*
3. *$(C_+^\times, A_+^\times)$ and $(C_-^\times, A_-^\times)$ are observable,*
4. *$(A_+^\times, B_+^\times)$ and $(A_-^\times, B_-^\times)$ are exactly controllable,*
5. *$A_+$, $A_-$, $A_+^\times$ and $A_-^\times$ are all strongly bi-stable, and finally*
6. *$W(z)$ and $W(z)^{-1}$ have realizations of the form*

$$W(z) = zC_+(I_{\mathcal{X}_+} - zA_+)^{-1}B_+ + D + C_-(zI_{\mathcal{X}_-} - A_-)^{-1}B_- \tag{2.18}$$
$$W(z)^{-1} = zC_+^\times(I_{\mathcal{X}_+^\times} - zA_+^\times)^{-1}B_+^\times + D^\times + C_-^\times(zI_{\mathcal{X}_-^\times} - A_-^\times)^{-1}B_-^\times. \tag{2.19}$$

We emphasize that the meaning of (2.18) and (2.19) is

$$W(z) = \sum_{n\in\mathbb{Z}} W_n z^n \text{ where } W_n = \begin{cases} D & \text{if } n = 0, \\ C_+(A_+)^{n-1}B_+ & \text{if } n \geq 1, \text{ and} \\ C_-(A_-)^{(-n-1)}B_- & \text{if } n \leq -1 \end{cases}$$

$$W(z)^{-1} = \sum_{n\in\mathbb{Z}} [W^{-1}]_n z^n \text{ where } [W^{-1}]_n = \begin{cases} D^\times & \text{if } n = 0, \\ C_+^\times(A_+^\times)^{n-1}B_+^\times & \text{if } n \geq 1, \text{ and} \\ C_-^\times(A_-^\times)^{(-n-1)}B_-^\times & \text{if } n \leq -1. \end{cases}$$

Assumption (1) in Definition (2.2) implies that $\sum_{n=1}^\infty C_+A_+^{n-1}z^n$ is in $H^2_{\mathcal{L}(\mathcal{X}_+,\mathcal{U})}$ and that $\sum_{n=-\infty}^{-1} C_-(A_-)^{(-n-1)}z^n$ is in $H^{2\perp}_{\mathcal{L}(\mathcal{X}_-,\mathcal{U})}$ and hence $W \in L^2_{\mathcal{L}(\mathcal{U})}$. Similarly, assumption (4) in Definition (2.2) implies that $\sum_{n=1}^\infty (A_+^\times)^{n-1}B_+^\times z^n$ is in $H^2_{\mathcal{L}(\mathcal{U},\mathcal{X}_+^\times)}$ and $\sum_{n=-\infty}^{-1}(A_-^\times)^{(-n-1)}B_-^\times z^n$ is in $H^{2\perp}_{\mathcal{L}(\mathcal{U},\mathcal{X}_-^\times)}$ and hence $W^{-1} \in L^2_{\mathcal{L}(\mathcal{U})}$. Thus *both $W$ and $W^{-1}$ are in $L^2_{\mathcal{L}(\mathcal{U})}$ whenever $W$ has a strongly bi-dichotomous realization.*

## 3. Dual shift-invariant pair of subspaces via strongly regular Beurling-Lax representer

We say that the subspace $\mathcal{M}$ of $L^2_{\mathcal{U}}$ is *shift-invariant* if $\mathcal{M}$ is invariant for the shift operator $M_z\colon f(z) \mapsto zf(z)$. Such a subspace is said to be *simply-invariant* if

$$\bigcap_{n=0}^{\infty} (M_z)^n \mathcal{M} = \{0\}$$

and to be *full-range* if

$$\bigcup_{n=0}^{\infty} (M_{z^{-1}})^n \mathcal{M} \text{ is dense in } L^2_{\mathcal{U}}.$$

Similarly, a subspace $\mathcal{M}^\times$ is said to be *backward-shift-invariant* if $M_{z^{-1}}\mathcal{M}^\times \subset \mathcal{M}^\times$. Such a subspace is *simply-invariant* (for the backward shift) if in addition

$$\bigcap_{n=0}^{\infty} (M_{z^{-1}})^n \mathcal{M}^\times = \{0\}$$

and to be *full-range* if

$$\bigcup_{n=0}^{\infty} (M_z)^n \mathcal{M}^\times \text{ is dense in } L^2_{\mathcal{U}}.$$

As a short hand we define a *dual shift-invariant pair* of subspaces as follows.

**Definition 3.1.** *Let $\mathcal{M}$ and $\mathcal{M}^\times$ be a pair of (closed) subspaces of $L^2_{\mathcal{U}}$. We say that the pair $(\mathcal{M}, \mathcal{M}^\times)$ is a* dual shift-invariant pair *if*

1. *$\mathcal{M}$ is full-range simply-invariant for the forward shift $M_z$,*
2. *$\mathcal{M}^\times$ is full-range simply-invariant for the backward shift $M_{z^{-1}}$, and*
3. *$\mathcal{M}$ and $\mathcal{M}^\times$ form a direct sum decomposition of the whole space:*

$$\mathcal{M} \dot{+} \mathcal{M}^\times = L^2_{\mathcal{U}}. \tag{3.1}$$

*Remark* 3.2. In Definition 3.1, the conditions (1) and (2) can be weakened either to

1′ $\mathcal{M}$ is simply-invariant for $M_z$,
2′ $\mathcal{M}^\times$ is simply-invariant for $M_{z^{-1}}$

or to

1″ $\mathcal{M}$ is full-range for $M_z$,
2″ $\mathcal{M}^\times$ is full-range for $M_{z^{-1}}$.

Indeed, given that $\mathcal{M}$ is invariant for $M_z$ and that $\mathcal{M}^\times$ is invariant for $M_{z^{-1}}$, then application of the bounded invertible operator $M_{z^N}$ (respectively $M_{z^{-N}}$) to the decomposition $\mathcal{M}^\times \dot{+} \mathcal{M}$ gives

$$L^2_{\mathcal{U}} = z^N \mathcal{M}^\times \dot{+} z^N \mathcal{M}, \qquad L^2_{\mathcal{U}} = z^{-N} \mathcal{M}^\times \dot{+} z^{-N} \mathcal{M}$$

for all $N = 1, 2, \dots$. From the first decomposition, we see that then $\mathcal{M}^\times$ is full-range for $M_{z^{-1}}$ if and only if $\mathcal{M}$ is simply-invariant for $M_z$. From the second decomposition, we see that $\mathcal{M}^\times$ is simply-invariant for $M_{z^{-1}}$ if and only if $\mathcal{M}$ is full-range for $M_z$.

**Definition 3.3.** *We say that the $\mathcal{L}(\mathcal{U})$-valued function $W$ on the unit circle is $L^2$-* regular *if*

1. $W(\zeta)^{-1}$ *exists for almost all $\zeta \in \mathbb{T}$ and both $W$ and $W^{-1}$ are in $L^2_{\mathcal{L}(\mathcal{U})}$, i.e., multiplication by either $W$ or $W^{-1}$ maps $\mathcal{U}$ into $L^2_{\mathcal{U}}$, and*
2. *the operator*
$$M_W P_{H^2_{\mathcal{U}}} M_W^{-1} \colon L^\infty_{\mathcal{U}}(\mathbb{T}) \to L^1_{\mathcal{U}}(\mathbb{T})$$
*has range in $L^2_{\mathcal{U}}(\mathbb{T})$ and extends to define a bounded operator from $L^2_{\mathcal{U}}$ into itself.*

The next result from [21] gives the fundamental connection between these objects. We include a more polished version of the proof from [21] which handles the case dim $\mathcal{U} = \infty$.

**Theorem 3.4.** *(See* [21]*.) Suppose that $(\mathcal{M}, \mathcal{M}^\times)$ are two subspaces of $L^2_{\mathcal{U}}$. Then the following are equivalent:*

1. *$(\mathcal{M}, \mathcal{M}^\times)$ is a dual shift-invariant pair.*
2. *There exists a $L^2$-regular $\mathcal{L}(\mathcal{U})$-valued function $W$ so that*
$$\mathcal{M} = L^2\text{-}clos\ W H^\infty_{\mathcal{U}} \text{ and } \mathcal{M}^\times = L^2\text{-}clos\ W \overline{H^\infty_{\mathcal{U},0}}. \tag{3.2}$$

*Moreover, in this case the dual shift-invariant pair $(\mathcal{M}, \mathcal{M}^\times)$ uniquely determines $W$ up to an invertible constant right factor, i.e., if $W'$ is another $L^2$-regular $\mathcal{L}(\mathcal{U})$-valued function as in part (2), then there is an invertible constant operator $X \in \mathcal{L}(\mathcal{U})$ so that $W'(z) = W(z)X$.*

*Proof of (1) $\implies$ (2):* Given a subspace of $\mathcal{N}$ of $L^2_{\mathcal{U}}$, we shall abuse notation and write simply $z^N\mathcal{N}$ for the space $M_{z^N}\mathcal{N} = z^n \cdot \mathcal{N}$. From the assumption
$$L^2_{\mathcal{U}} = \mathcal{M}^\times \dot{+} \mathcal{M}, \tag{3.3}$$
we also have
$$L^2_{\mathcal{U}} = zL^2_{\mathcal{U}} = z\mathcal{M}^\times \dot{+} z\mathcal{M}. \tag{3.4}$$
Since $\mathcal{M}^\times \subset z\mathcal{M}^\times$ and $z\mathcal{M} \subset \mathcal{M}$, combining (3.3) and (3.4) gives
$$L^2_{\mathcal{U}} = \mathcal{M}^\times \dot{+} \mathcal{L} \dot{+} z\mathcal{M}, \qquad \mathcal{M} = \mathcal{L} \dot{+} z\mathcal{M} \tag{3.5}$$
where we have set
$$\mathcal{L} = z\mathcal{M}^\times \cap \mathcal{M}. \tag{3.6}$$
Similarly, from
$$L^2_{\mathcal{U}} = z^{-1}L^2_{\mathcal{U}} = z^{-1}\mathcal{M}^\times \dot{+} z^{-1}\mathcal{M}$$
we get
$$L^2_{\mathcal{U}} = z^{-1}\mathcal{M}^\times \dot{+} z^{-1}\mathcal{L} \dot{+} \mathcal{M}, \qquad \mathcal{M}^\times = z^{-1}\mathcal{M}^\times \dot{+} z^{-1}\mathcal{L}.$$

An inductive argument then gives

$$\begin{aligned}
L^2_{\mathcal{U}} &= z^{-N}\mathcal{M}^\times \dot{+} z^{-N}\mathcal{L} \dot{+} \cdots \dot{+} z^{-1}\mathcal{L} \dot{+} \mathcal{L} \dot{+} z\mathcal{L} \dot{+} \cdots \dot{+} z^{N-1}\mathcal{L} \dot{+} z^N\mathcal{M}, \\
\mathcal{M} &= \mathcal{L} \dot{+} z\mathcal{L} \dot{+} \cdots \dot{+} z^{N-1} \cdot \mathcal{L} \dot{+} z^N\mathcal{M}, \\
\mathcal{M}^\times &= z^{-N}\mathcal{M}^\times \dot{+} z^{-N}\mathcal{L} \dot{+} \cdots \dot{+} z^{-1}\mathcal{L} \qquad (3.7)
\end{aligned}$$

for $N = 0, 1, 2, \dots$ (with the interpretation $L^2_{\mathcal{U}} = \mathcal{M}^\times \dot{+} \mathcal{M}$, $\mathcal{M} = \mathcal{M}$ and $\mathcal{M}^\times = \mathcal{M}^\times$ for the $N = 0$ case). Note that the projection onto $z^N\mathcal{M}$ along the subspace

$$z^{-N}\mathcal{M}^\times \dot{+} \left[\dot{+}_{n=-N}^{N-1} z^n\mathcal{L}\right] = z^N\mathcal{M}^\times$$

in the first line of (3.7) is given by

$$Q_N := M_{z^N} P_{\mathcal{M}} M_{z^{-N}} \qquad (3.8)$$

where $P_{\mathcal{M}}$ denotes the projection onto $\mathcal{M}$ along $\mathcal{M}^\times$. The estimate

$$\|Q_N\| = \|M_{z^N} P_{\mathcal{M}} M_{z^{-N}}\| \leq \|P_{\mathcal{M}}\|$$

then shows that $Q_N$ is uniformly bounded in $N$ and hence a subsequence converges to an operator $Q_\infty$ in the weak-$*$ topology. From the fact that $Q_{N_0}Q_N = Q_{N_0}$ once $N > N_0$, we see that actually the whole sequence $\{Q_N\}_{N=1,2,\dots}$ converges to $Q_\infty$ as $N \to \infty$ and that $Q_\infty$ is a projection ($Q^2_\infty = Q_\infty$). One can check that $Q_\infty$ is a projection of $L^2_{\mathcal{U}}$ onto $\cap_{N\geq 0} z^N\mathcal{M}$ along $L^2$-clos $\cup_{N\geq 0} z^N\mathcal{M}^\times$. The hypothesis that $\mathcal{M}$ is simply invariant for $M_z$ implies that $Q_\infty = 0$, and hence $P_N = I - Q_N$ converges to the identity operator $I_{L^2_{\mathcal{U}}}$. From the decomposition for $z^N\mathcal{M}^\times$

$$z^N\mathcal{M}^\times = \mathcal{M}^\times \dot{+} \left[\dot{+}_{n=0}^{N} z^n\mathcal{L}\right]$$

we see that, in particular, each $f \in \mathcal{M}$ has the generalized Fourier series expansion (with limit in the weak-$L^2_{\mathcal{U}}$ topology)

$$f = \lim_{N\to\infty} \sum_{n=0}^{N} z^n \ell_n \text{ with } \ell_n = P_{\mathcal{L}} M_{z^{-n}} f \text{ for all } n \in \mathbb{Z}_+ \text{ for all } f \in \mathcal{M}. \qquad (3.9)$$

Similarly, given that $\mathcal{M}^\times$ is simply invariant for $M_{z^{-1}}$, it follows that elements of $\mathcal{M}^\times$ have a Fourier series expansion

$$f = \text{w-lim}_{N\to\infty} \sum_{n=-1}^{-N} z^n \ell_n \text{ with } \ell_n = P_{\mathcal{L}} M_{z^{-n}} f \text{ for all } n \in \mathbb{Z}_- \text{ for all } f \in \mathcal{M}^\times. \qquad (3.10)$$

Hence elements $f$ of $L^2$ have a bilateral Fourier series expansion

$$f = \text{w-lim}_{N\to\infty} \sum_{n=-N}^{N} z^n \ell_n \text{ with } \ell_n = P_{\mathcal{L}} M_{z^{-n}} f \text{ for all } n \in \mathbb{Z} \text{ for all } f \in L^2_{\mathcal{U}}. \qquad (3.11)$$

While it need not be the case that the series (3.11) converges for an arbitrary norm-square summable sequence $\{\ell_n\}_{n\in\mathbb{Z}}$, it is easily verified that absolute summability ($\sum_{n\in\mathbb{Z}} \|\ell_n\| < \infty$) is sufficient for (even norm) convergence of the series (3.11):

$$\sum_{n\in\mathbb{Z}} \|\ell_n\| < \infty \Longrightarrow \lim_{N\to\infty} \sum_{n=-N}^{N} z^n \ell_n \text{ exists.} \tag{3.12}$$

Choose $\widetilde{\mathcal{U}}$ to be a Hilbert space of the same dimension as $\mathcal{L}$ and let $W\colon \widetilde{\mathcal{U}} \to \mathcal{L}$ be an isomorphism. Then we can view $W$ as an operator-valued function in $L^2_{\mathcal{L}(\widetilde{\mathcal{U}},\mathcal{U})}$. The Fourier decompositions (3.9), (3.10) and (3.11) tell us that

$$W\cdot \mathcal{P}_{\widetilde{\mathcal{U}}} \text{ is dense in } \mathcal{M}, \qquad W\cdot \overline{\mathcal{P}_{\widetilde{\mathcal{U}},0}} \text{ is dense in } \mathcal{M}^{\times}, \qquad W\cdot \mathcal{T}_{\widetilde{\mathcal{U}}} \text{ is dense in } L^2_{\mathcal{U}} \tag{3.13}$$

where $\mathcal{P}_{\widetilde{\mathcal{U}}}$ denotes the space of analytic polynomials $\sum_{n=0}^{N} p_n z^n$ ($N < \infty$) with coefficients $p_n \in \mathcal{U}$, $\overline{\mathcal{P}_{0,\widetilde{\mathcal{U}}}}$ denotes the space of strictly antianalytic polynomials $\sum_{n=1}^{N} p_n z^{-n}$ with coefficients $p_n \in \widetilde{\mathcal{U}}$, and $\mathcal{T}_{\widetilde{\mathcal{U}}}$ denotes the space of all trigonometric polynomials with coefficients in $\widetilde{\mathcal{U}}$.

Note that if $(\mathcal{M}, \mathcal{M}^{\times})$ is a dual shift-invariant pair with $P_{\mathcal{M}}$ equal to the projection onto $\mathcal{M}$ along $\mathcal{M}^{\times}$, then $((\mathcal{M}^{\times})^{\perp}, \mathcal{M}^{\perp})$ is also a dual shift-invariant pair with $(P_{\mathcal{M}})^*$ equal to the projection onto $(\mathcal{M}^{\times})^{\perp}$ along $\mathcal{M}^{\perp}$. We apply the preceding analysis to produce an $L^2_{\mathcal{L}(\mathcal{U})}$-function $W_{\perp}$ and a second coefficient Hilbert space $\widetilde{\mathcal{U}}_{\perp}$ so that

$$W_{\perp}\widetilde{\mathcal{U}}_{\perp} = z\mathcal{M}^{\perp} \cap (\mathcal{M}^{\times})^{\perp} = (z\mathcal{M})^{\perp} \cap (\mathcal{M}^{\times})^{\perp} =: \mathcal{L}_{\perp}. \tag{3.14}$$

Thus, since $W\mathcal{P}_{\mathcal{U},0} \subset z\mathcal{M}$ and $W\overline{\mathcal{P}_{\mathcal{U},0}} \subset \mathcal{M}^{\times}$, for each $u \in \mathcal{U}$ and $u_{\perp} \in \widetilde{\mathcal{U}}_{\perp}$,

$$\langle W_{\perp}u_{\perp}, Wpu\rangle_{L^2_{\mathcal{U}}} = 0 \text{ for all } p \in \mathcal{P}_0 + \overline{\mathcal{P}_0}.$$

It follows that the $L^1_{\mathcal{L}(\widetilde{\mathcal{U}},\widetilde{\mathcal{U}}_{\perp})}$ function $W_{\perp}^*W$ has all its Fourier coefficients $[W_{\perp}^*W]_n$ equal to zero with the exception of the zero coefficient $[W_{\perp}^*W]_0$. It follows that the operator-valued function $W_{\perp}^*W$ is equal to a constant, say $L \in \mathcal{L}(\widetilde{\mathcal{U}},\widetilde{\mathcal{U}}_{\perp})$.

We argue next that this constant $L$ is invertible. To see this, note from the definition (3.14) of $\mathcal{L}_{\perp}$ that the orthogonal complement of $\mathcal{L}_{\perp}$ is given by

$$(\mathcal{L}_{\perp})^{\perp} = \mathcal{M}^{\times} \dot{+} z\mathcal{M}.$$

From the direct-sum decomposition (3.5) we conclude that

$$L^2_{\mathcal{U}} = \mathcal{L} \dot{+} (\mathcal{L}_{\perp})^{\perp}.$$

Thus, if $\mathbf{P}_{\mathcal{L}_{\perp}}$ is the orthogonal projection of $L^2_{\mathcal{U}}$ onto $\mathcal{L}_{\perp}$, we have

$$\mathcal{L}_{\perp} = \text{Ran } \mathbf{P}_{\mathcal{L}_{\perp}} = \text{Ran}\,(\mathbf{P}_{\mathcal{L}_{\perp}}|_{\mathcal{L}}), \qquad \text{Ker}\,(\mathbf{P}_{\mathcal{L}_{\perp}}|_{\mathcal{L}}) = \mathcal{L} \cap (\mathcal{L}_{\perp})^{\perp} = \{0\}$$

and it follows from the open mapping theorem (see e.g. [41, Corollaries 2.12]) that the operator

$$\mathbf{P}_{\mathcal{L}_{\perp}}|_{\mathcal{L}}\colon \mathcal{L} \to \mathcal{L}_{\perp}$$

is bounded and invertible. By construction we have that the restricted multiplication operators

$$M_W|_{\widetilde{\mathcal{U}}} = \mathbf{P}_{\mathcal{L}} M_W|_{\widetilde{\mathcal{U}}} : \widetilde{\mathcal{U}} \to \mathcal{L}, \qquad M_{W_\perp}|_{\widetilde{\mathcal{U}}_\perp} = \mathbf{P}_{\mathcal{L}_\perp} M_{W_\perp}|_{\widetilde{\mathcal{U}}_\perp} : \widetilde{\mathcal{U}}_\perp \to \mathcal{L}_\perp$$

are bounded and invertible and we can write

$$\mathbf{P}_{\widetilde{\mathcal{U}}_\perp}(M_{W_\perp})^* = \mathbf{P}_{\widetilde{\mathcal{U}}_\perp}(M_{W_\perp})^* \mathbf{P}_{\mathcal{L}_\perp}$$

where $\mathbf{P}_{\widetilde{\mathcal{U}}_\perp}$ and $\mathbf{P}_{\mathcal{L}_\perp}$ are the orthogonal projections from $L^2_{\widetilde{\mathcal{U}}_\perp}$ onto $\widetilde{\mathcal{U}}_\perp$ and from $L^2_{\mathcal{U}}$ onto $\mathcal{L}_\perp$ respectively. Hence we have

$$\begin{aligned} L = (M_{W_\perp})^* M_W|_{\widetilde{\mathcal{U}}} &= \mathbf{P}_{\widetilde{\mathcal{U}}_\perp}(M_{W_\perp})^* M_W|_{\widetilde{\mathcal{U}}} = \mathbf{P}_{\widetilde{\mathcal{U}}_\perp}(M_{W_\perp})^* \mathbf{P}_{\mathcal{L}_\perp} M_W|_{\widetilde{\mathcal{U}}} \\ &= \left(\mathbf{P}_{\mathcal{L}_\perp} M_{W_\perp}|_{\widetilde{\mathcal{U}}_\perp}\right)^* \cdot (\mathbf{P}_{\mathcal{L}_\perp}|_{\mathcal{L}}) \cdot (\mathbf{P}_{\mathcal{L}} M_W|_{\widetilde{\mathcal{U}}}) \end{aligned}$$

is the composition of isomorphisms and hence is invertible.

A parallel analysis (not needed for the proof of the theorem) shows that

$$\mathcal{L}^\perp = \mathcal{M}^\perp \dot{+} z\mathcal{M}^{\times\perp}$$

from which we deduce that

$$L^2_{\mathcal{U}} = \mathcal{L}_\perp \dot{+} \mathcal{L}^\perp$$

and that the operator

$$\mathbf{P}_{\mathcal{L}}|_{\mathcal{L}_\perp} : \mathcal{L}_\perp \to \mathcal{L}$$

is invertible, and finally that

$$L^* = \left(\mathbf{P}_{\mathcal{L}} M_W|_{\widetilde{\mathcal{U}}}\right)^* \cdot (\mathbf{P}_{\mathcal{L}}|_{\mathcal{L}_\perp}) \cdot \left(\mathbf{P}_{\mathcal{L}_\perp} M_{W_\perp}|_{\widetilde{\mathcal{U}}_\perp}\right) : \widetilde{\mathcal{U}}_\perp \to \widetilde{\mathcal{U}}$$

is invertible; of course this last fact follows directly from the fact shown above that $\Gamma \in \mathcal{L}(\widetilde{\mathcal{U}}, \widetilde{\mathcal{U}}_\perp)$ is invertible by taking the adjoint.

Since

$$W\mathcal{P}_{\mathcal{U}} = \cup_{N>0}[z^{-N}\mathcal{L} \dot{+} \cdots \dot{+} z^{-1}\mathcal{L} \dot{+} \mathcal{L} \dot{+} z\mathcal{L} \dot{+} \cdots \dot{+} z^{N-1}\mathcal{L}]$$

is dense in $L^2_{\mathcal{U}}$ by the full-range/simply-invariant assumptions, we see that the range of $W(\zeta)$ must be dense for almost all $\zeta \in \mathbb{T}$. The previous analysis shows that $\Gamma^{-1}W_\perp(\zeta)^*$ serves as a left inverse for $W(\zeta)$ almost everywhere. We conclude that in fact $W(\zeta)$ is invertible for almost all $\zeta$ and without loss of generality we may take $\widetilde{\mathcal{U}} = \widetilde{\mathcal{U}}_\perp = \mathcal{U}$.

Since the operator $M_W P_{H^2_{\mathcal{U}}} M_{W^{-1}}$ represents the projection of $L^2_{\mathcal{U}}$ onto $\mathcal{M}$ along $\mathcal{M}^\times$ (which is bounded by hypothesis), we see that $M_W P_{H^2_{\mathcal{U}}} M_{W^{-1}}$ has a bounded extension to an operator defined on all of $L^2_{\mathcal{U}}$. It follows now that $W \in L^2_{\mathcal{L}(\mathcal{U})}$ meets all the requirements of statement (2) in the theorem. □

*Proof of (2) $\Longrightarrow$ (1):* The converse follows by essentially reversing the argument in the proof of (1) $\Longrightarrow$ (2). Suppose that $W$ is as in (2). Necessarily the bounded operator on $L^2_{\mathcal{U}}$ arising from the extension of $M_W P_{H^2_{\mathcal{U}}} M_{W^{-1}}$ acting on $L^\infty_{\mathcal{U}}$ is equal to a (bounded) projection operator $P_{\mathcal{M}}$ of $L^2_{\mathcal{U}}$ onto $\mathcal{M}$ along $\mathcal{M}^\times$. Thus (3.1) holds.

As in the proof of (1) $\Longrightarrow$ (2) we pick up the Fourier series decompositions (3.7) with $\mathcal{L} = z\mathcal{M}^{\times} \bigcap \mathcal{M}$. But from the assumption (3.2) we see also that

$$W\mathcal{U} \subset \mathcal{L}$$

and hence

$$L^2\text{-clos}\, W\mathcal{P}_{\mathcal{U}} = \text{Ran}\,(P_{\mathcal{M}} - Q_{\infty})$$

where $Q_{\infty} = \text{weak-}*\ \lim_{N\to\infty} Q_N$ with $Q_N$ as in (3.8). But on the other hand, since $W$ is a Beurling-Lax representer for $\mathcal{M}$ (see (3.2)), necessarily

$$L^2\text{-clos}\, W\mathcal{P}_{\mathcal{U}} = \mathcal{M} = \text{Ran}\ P_{\mathcal{M}}.$$

We conclude that $Q_{\infty} = 0$, i.e., $\mathcal{M}$ is simply invariant. That $\mathcal{M}^{\times}$ is simply invariant for $M_{z^{-1}}$ can be proved in a similar way. By Remark 3.2, it follows now that $(\mathcal{M}, \mathcal{M}^{\times})$ is a dual shift-invariant pair, and (1) follows as wanted.

Along the way we showed that any representer $W$ must arise via the construction given in the proof of (1) $\Longrightarrow$ (2), and hence is uniquely determined up to a choice of identification map between $\mathcal{U}$ and $\mathcal{L}$; this observation leads to the uniqueness statement in Theorem 3.4. □

## 4. Dual shift-invariant pair of subspaces via pole-zero data

**Definition 4.1.** *Let $(C, A, Z, B, \Gamma)$ be a collection of operators such that*

$$C\colon \mathcal{X}_{\mathcal{P}} \to \mathcal{U}, \quad A\colon \mathcal{X}_{\mathcal{P}} \to \mathcal{X}_{\mathcal{P}}, \quad Z\colon \mathcal{X}_{\mathcal{Z}} \to \mathcal{X}_{\mathcal{Z}}, \quad B\colon \mathcal{U} \to \mathcal{X}_{\mathcal{Z}}, \quad \Gamma\colon \mathcal{X}_{\mathcal{P}} \to \mathcal{X}_{\mathcal{Z}}.$$

*Then we say that $\mathfrak{S} := (C, A, Z, B, \Gamma)$ is an* admissible Sylvester data set *if*

1. *the output pair $(C, A)$ is exactly observable,*
2. *the input pair $(Z, B)$ is exactly controllable, and*
3. *the* coupling operator *$\Gamma$ is a (possibly unbounded) closed operator with dense domain $\mathcal{D}(\Gamma) \subset \mathcal{X}_{\mathcal{P}}$ such that $A\colon \mathcal{D}(\Gamma) \to \mathcal{D}(\Gamma)$ and $\Gamma$ satisfies the Sylvester equation*

$$\Gamma A x - Z \Gamma x = BCx \text{ for all } x \in \mathcal{D}(\Gamma). \tag{4.1}$$

Note that if $(C, A, Z, B, \Gamma)$ is an admissible Sylvester data set, then the discussion above concerning the Stein equations (2.5) and (2.11) tells us that $A$ and $Z^*$ are automatically strongly stable (i.e., (2.7) and (2.13) hold). Moreover, since the exact controllability of $(Z, B)$ implies that $(Z, B)$ is also input-stable, it follows that both functions $[(zI - Z)^{-1}B]$ and $[(I - zZ)^{-1}B]$ are in $L^2_{\mathcal{L}(\mathcal{X}_{\mathcal{Z}}, \mathcal{U})}$. As a consequence of Proposition 2.1, it then follows that, for any $h \in L^2_{\mathcal{U}}$, both $(zI - Z)^{-1}Bh(z)$ and $(I - zZ)^{-1}Bh(z)$ are in $L^1_{\mathcal{X}_{\mathcal{Z}}}$ and in particular have well-defined Fourier coefficients.

We associate with any admissible Sylvester data set a subspace of $L^2_{\mathcal{U}}$ in two distinct ways:

$$\begin{aligned} \mathcal{M}_{\mathfrak{S}} :=& \{C(zI-A)^{-1}x + f(z) \colon x \in \mathcal{D}(\Gamma) \text{ and } f \in H^2_{\mathcal{U}} \text{ such that} \\ & \big[(zI-Z)^{-1}Bf(z)\big]_{-1} = \Gamma x\}, \\ \mathcal{M}^{\times}_{\mathfrak{S}} :=& \{g(z) + C(I-zA)^{-1}y \colon g \in H^{2\perp}_{\mathcal{U}} \text{ and } y \in \mathcal{D}(\Gamma) \text{ such that} \\ & \big[(I-zZ)^{-1}Bg(z)\big]_{-1} = \Gamma y\}. \end{aligned} \tag{4.2}$$

Here $[h(z)]_{-1}$ is the coefficient of $z^{-1}$ in the Fourier series $h(z) \sim \sum_{n=-\infty}^{\infty} h(n)z^n$ for a function $h \in L^1_{\mathcal{X}_{\mathcal{Z}}}$. In terms of control and observation operators, we can write instead

$$\mathcal{M}_{\mathfrak{S}} := \{\widehat{\mathcal{O}}^b_{C,A}x + f \colon x \in \mathcal{D}(\Gamma) \text{ and } f \in H^2_{\mathcal{U}} \text{ such that } \widehat{\mathcal{C}}^b_{Z,B}f = \Gamma x\}, \tag{4.3}$$

$$\mathcal{M}^{\times}_{\mathfrak{S}} := \{g + \widehat{\mathcal{O}}^f_{C,A}y \colon g \in H^{2\perp}_{\mathcal{U}} \text{ and } y \in \mathcal{D}(\Gamma) \text{ such that } \widehat{\mathcal{C}}^f_{Z,B}g = \Gamma y\}. \tag{4.4}$$

Given a pair

$$\mathfrak{S} = (C, A, Z, B, \Gamma), \qquad \mathfrak{S}^{\times} = (C^{\times}, A^{\times}, Z^{\times}, B^{\times}, \Gamma^{\times}) \tag{4.5}$$

of Sylvester data sets, we define the *coupling matrix* $\Gamma_{\mathfrak{S},\mathfrak{S}^{\times}}$ by

$$\Gamma_{\mathfrak{S},\mathfrak{S}^{\times}} = \begin{bmatrix} -\widehat{\mathcal{C}}^f_{Z^{\times},B^{\times}}\widehat{\mathcal{O}}^b_{C,A} & -\Gamma^{\times} \\ \Gamma & \widehat{\mathcal{C}}^b_{Z,B}\widehat{\mathcal{O}}^f_{C^{\times},A^{\times}} \end{bmatrix} \colon \begin{bmatrix} \mathcal{D}(\Gamma) \\ \mathcal{D}(\Gamma^{\times}) \end{bmatrix} \to \begin{bmatrix} \mathcal{X}^{\times}_{\mathcal{Z}} \\ \mathcal{X}_{\mathcal{Z}} \end{bmatrix}. \tag{4.6}$$

Then we have the following result.

**Theorem 4.2.** 1. *$\mathcal{M}$ is a closed shift-invariant subspace of $L^2_{\mathcal{U}}$ if and only if $\mathcal{M}$ has the form $\mathcal{M}_{\mathfrak{S}}$ for an admissible Sylvester data set*

$$\mathfrak{S} = (C, A, Z, B, \Gamma).$$

*Moreover,*

(a) *$\mathcal{M}$ is simply invariant for $M_z$ if and only if in addition $A^*$ is strongly stable (so $A$ is strongly bi-stable), and*

(b) *$\mathcal{M}$ is full range if and only if $Z$ is strongly stable (so $Z$ is strongly bi-stable).*

2. *$\mathcal{M}^{\times}$ is a closed backward-shift-invariant subspace of $L^2_{\mathcal{U}}$ if and only if $\mathcal{M}^{\times}$ has the form $\mathcal{M}^{\times} = \mathcal{M}^{\times}_{\mathfrak{S}^{\times}}$ for an admissible Sylvester data set*

$$\mathfrak{S}^{\times} = (C^{\times}, A^{\times}, Z^{\times}, B^{\times}, \Gamma^{\times}). \tag{4.7}$$

*Moreover*

(a) *$\mathcal{M}^{\times}$ is simply invariant for $M_{z^{-1}}$ if and only if in addition $(A^{\times})^*$ is strongly stable (so $A^{\times}$ is strongly bi-stable).*

(b) *$\mathcal{M}^{\times}$ is full range if and only if $Z^{\times}$ is strongly stable (so $Z^{\times}$ is strongly bi-stable).*

3. *Suppose that* $(\mathcal{M}, \mathcal{M}^\times)$ *is a pair of subspaces of* $L^2_{\mathcal{U}}$ *of the form* $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ *and* $\mathcal{M}^\times = \mathcal{M}^\times_{\mathfrak{S}^\times}$ *(so* $\mathcal{M}$ *is closed and shift-invariant and* $\mathcal{M}^\times$ *is closed and backward-shift-invariant). Then* $(\mathcal{M}, \mathcal{M}^\times)$ *satisfies the matching condition*
$$L^2_{\mathcal{U}} = \mathcal{M}^\times \dot{+} \mathcal{M} \tag{4.8}$$
*if and only if the coupling matrix*
$$\Gamma_{\mathfrak{S},\mathfrak{S}^\times} \colon \begin{bmatrix} \mathcal{D}(\Gamma) \\ \mathcal{D}(\Gamma^\times) \end{bmatrix} \to \begin{bmatrix} \mathcal{X}^\times_{\mathcal{Z}} \\ \mathcal{X}_{\mathcal{Z}} \end{bmatrix} \tag{4.9}$$
*associated with* $\mathfrak{S}$ *and* $\mathfrak{S}^\times$ *as in* (4.6) *is invertible. More precisely,*
$$\dim(\mathcal{M} \cap \mathcal{M}^\times) = \dim \operatorname{Ker} \Gamma_{\mathfrak{S},\mathfrak{S}^\times}, \tag{4.10}$$
$$\dim L^2_{\mathcal{U}}/(\mathcal{M} + \mathcal{M}^\times) = \dim(\mathcal{X}_{\mathcal{Z}} \oplus \mathcal{X}^\times_{\mathcal{Z}})/\operatorname{Ran} \Gamma_{\mathfrak{S},\mathfrak{S}^\times}. \tag{4.11}$$

*Proof of (1):* Suppose first that $\mathcal{M}$ has the form $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ for an admissible Sylvester data set $\mathfrak{S}$ as in (4.7). One can see that $\mathcal{M}$ is a closed subspace of $L^2_{\mathcal{U}}$ as follows. Suppose that $\mathcal{O}^b_{C,A} x_n + f_n \in \mathcal{M}$ and converges to $k + f$ in $L^2_{\mathcal{U}}$ where $k \in H^{2\perp}_{\mathcal{U}}$ and $f \in H^2_{\mathcal{U}}$. Then necessarily
$$\widehat{\mathcal{O}}^b_{C,A} x_n \to k \in H^{2\perp}_{\mathcal{U}} \text{ and } f_n \to f \in H^2_{\mathcal{U}} \text{ as } n \to \infty$$
where we know that
$$\widehat{\mathcal{C}}^b_{Z,B} f_n = \Gamma x_n \text{ for all } n = 1, 2, \dots .$$
Since $(C, A)$ is exactly observable, from the fact that $\mathcal{O}^b_{C,A} x_n$ converges to $k$ it follows that $x_n$ converges to some $x \in \mathcal{X}_{\mathcal{P}}$ and then $k = \mathcal{O}^b_{C,A} x$. From the fact that $f_n \to f$ as $n \to \infty$ and that $\widehat{\mathcal{C}}^b_{Z,B}$ is bounded, we see that
$$\Gamma x_n = \widehat{\mathcal{C}}^b_{Z,B} f_n \to \widehat{\mathcal{C}}^b_{Z,B} f \text{ as } n \to \infty.$$
As $\Gamma$ is closed, we conclude that $x \in \mathcal{D}(\Gamma)$ and that $\Gamma x = \widehat{\mathcal{C}}^b_{Z,B} f$. We now see that $k + f = \widehat{\mathcal{O}}^b_{C,A} x + f \in \mathcal{M}$, and it follows that $\mathcal{M}$ is closed as claimed.

We argue that $\mathcal{M}_{\mathfrak{S}}$ is $M_z$-invariant as follows. Let $C(zI - A)^{-1}x + f \in \mathcal{M}_{\mathfrak{S}} = \mathcal{M}$, with $x \in \mathcal{D}(\Gamma)$ and $f \in H^2_{\mathcal{U}}$. Then
$$\begin{aligned} z[C(zI - A)^{-1}x + f] &= C(zI - A)^{-1}[A + (zI - A)]x + zf \\ &= C(zI - A)^{-1}Ax + (Cx + zf). \end{aligned}$$
while
$$\begin{aligned} [(zI - Z)^{-1}B(Cx + zf)]_{-1} &= BCx + \left[(zI - Z)^{-1}[Z + (zI - Z)]Bf\right]_{-1} \\ &= BCx + \left[(zI - Z)^{-1}ZBf + Bf\right]_{-1} \\ &= BCx + Z\left[(zI - Z)^{-1}Bf\right]_{-1} \\ &= BCx + Z\Gamma x = \Gamma A x \end{aligned}$$
by the Sylvester equation. Therefore $z[C(zI - A)^{-1}x + f] \in \mathcal{M}$, i.e., $\mathcal{M}$ is $M_z$-invariant.

Suppose conversely that $\mathcal{M}$ is a closed shift-invariant subspace of $L^2_{\mathcal{U}}$. Let $\mathcal{P} = L^2$-clos $P_{H^{2\perp}_{\mathcal{U}}}\mathcal{M}$ and $\mathcal{Z} = H^2_{\mathcal{U}} \ominus (\mathcal{M} \cap H^2_{\mathcal{U}})$. Then there is a closed operator $\Gamma\colon \mathcal{P} \to \mathcal{Z}$ with domain $\mathcal{D}(\Gamma) = P_{H^{2\perp}_{\mathcal{U}}}\mathcal{M}$ so that $\mathcal{M} = \{k + \Gamma k\colon k \in \mathcal{D}(\Gamma)\} \oplus (\mathcal{M} \cap H^2_{\mathcal{U}})$. Define $A\colon \mathcal{P} \to \mathcal{P}$ and $Z\colon \mathcal{Z} \to \mathcal{Z}$ by $A = P_{\mathcal{P}} M_z|_{\mathcal{P}}$ and $Z = P_{\mathcal{Z}} M_z|_{\mathcal{Z}}$. Also, define $C\colon \mathcal{P} \to \mathcal{U}$ and $B\colon \mathcal{U} \to \mathcal{Z}$ by $C(k) = [k]_{-1}$ and $B(f) = P_{\mathcal{Z}} f$. We claim that $\mathfrak{S} = (C, A, Z, B, \Gamma)$ is an admissible Sylvester data set and that $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$.

We first show that $(C, A)$ is output stable and exactly observable. Let $x = \sum_{k=-\infty}^{-1} u_k z^k \in \mathcal{P}$. Then $CA^n x = [P_{\mathcal{P}} M_{z^n} x]_{-1} = \left[P_{\mathcal{P}} \sum_{k=-\infty}^{-1} u_k z^{n+k}\right]_{-1} = u_{-n-1}$. Hence $\mathcal{O}^f_{C,A}\colon x \to \{CA^n x\}_{n=0,1,2,\ldots}$ maps $\mathcal{P}$ into $\ell^2_{\mathcal{U}}(\mathbb{Z}_+)$, so $(C, A)$ is output-stable. Since $A^* = P_{\mathcal{P}} M_{z^{-1}}|_{\mathcal{P}}$ and $C^* u = u z^{-1}$, we find that

$$\begin{aligned}\sum_{n=0}^{\infty} (A^*)^n C^* C A^n x &= \sum_{n=0}^{\infty} (A^*)^n C^* u_{-n-1} = \sum_{n=0}^{\infty} (A^*)^n u_{-n-1} z^{-1} \\ &= \sum_{n=0}^{\infty} u_{-n-1} z^{-n-1} = x.\end{aligned}$$

Thus $(\mathcal{O}^f_{C,A})^* \mathcal{O}^f_{C,A} = I$, so the observation operator is an isometry and $(C, A)$ is exactly observable as wanted.

We next check that $(Z, B)$ is input-stable and exactly controllable. If the sequence $\{u_k\}_{k \in \mathbb{Z}_+}$ is in $\ell^2_{\mathcal{U}}(\mathbb{Z}_+)$, then

$$\| \begin{bmatrix} B & ZB & Z^2 B & \cdots \end{bmatrix} \{u_k\}_{k \in \mathbb{Z}_+} \|^2_{\mathcal{X}} = \| \sum_{n=0}^{\infty} u_n z^n \|^2_{H^2_{\mathcal{U}}} = \| \{u_k\}_{k \in \mathbb{Z}_+} \|^2_{\ell_{\mathcal{U}}(\mathbb{Z}_+)}.$$

So $\begin{bmatrix} B & ZB & Z^2 B & \cdots \end{bmatrix}$ is bounded, i.e., $(Z, B)$ is input-stable. Also, since $Z^* = P_{\mathcal{Z}} M_{z^{-1}}|_{\mathcal{Z}}$ and $B^*\colon f \to f_0$, it follows readily that the control operator $\widehat{\mathcal{C}}^b_{Z,B}$ is an isometry, and thus $(Z, B)$ is exactly controllable.

The fact that $\Gamma$ is a closed operator follows from the fact that $\mathcal{M}$ is a closed subspace of $L^2_{\mathcal{U}}$. We next check the validity of the Sylvester equation (4.1). Let $k + \Gamma k + f \in \mathcal{M}$, where $k \in \mathcal{D}(\Gamma) = P_{H^{2\perp}_{\mathcal{U}}}\mathcal{M}$ and $f \in H^2_{\mathcal{U}} \cap \mathcal{M}$. Since $\mathcal{M}$ is shift-invariant, we may express $z(k + \Gamma k + f) = k' + \Gamma k' + f'$ for some $k' \in \mathcal{D}(\Gamma)$ and $f' \in H^2_{\mathcal{U}} \cap \mathcal{M}$. It follows that $k' = P_{H^{2\perp}_{\mathcal{U}}} zk$, $\Gamma k' = P_{\mathcal{Z}} k_{-1} + P_{\mathcal{Z}} M_z \Gamma k$, and $f' = P_{H^2_{\mathcal{U}} \cap \mathcal{M}} M_z \Gamma k + zf$. We have $BCk = Bk_{-1} = P_{\mathcal{Z}} k_{-1}$, and $Z\Gamma k = P_{\mathcal{Z}} M_z \Gamma k = \Gamma k' - P_{\mathcal{Z}} k_{-1}$. Therefore $\Gamma k' = (BC + Z\Gamma)k$. But $k' = P_{H^{2\perp}_{\mathcal{U}}} zk = P_{H^{2\perp}_{\mathcal{U}}} M_z|_{\mathcal{P}} k = P_{\mathcal{P}} M_z|_{\mathcal{P}} k = Ak$. Thus $\Gamma A k = BCk + Z\Gamma k$ for each $k \in \mathcal{D}(\Gamma)$, so the Sylvester equation is satisfied. We are now able to conclude that $\mathfrak{S} = (C, A, Z, B, \Gamma)$ an admissible Sylvester data set.

If $f(z) = \sum_{m=0}^{\infty} f_m z^m \in H^2_{\mathcal{U}}$, then

$$\begin{aligned}
[(zI - Z)^{-1} B f]_{-1} &= \left[ \sum_{n=0}^{\infty} Z^n z^{-n-1} B \left( \sum_{m=0}^{\infty} f_m z^m \right) \right]_{-1} \\
&= \sum_{n=0}^{\infty} P_{\mathcal{Z}} M_{z^n} f_n = P_{\mathcal{Z}} \sum_{n=0}^{\infty} f_n z^n = P_{\mathcal{Z}} f.
\end{aligned}$$

Also, if $x = \sum_{n=1}^{\infty} x_{-n} z^{-n} \in \mathcal{P}$, then $CA^n \colon x \to x_{-n-1}$ and hence

$$C(zI - A)^{-1} x = \sum_{n=0}^{\infty} (CA^n x) z^{-n-1} = \sum_{n=0}^{\infty} x_{-n-1} z^{-n-1} = x.$$

Hence

$$\begin{aligned}
\mathcal{M}_{\mathfrak{S}} &= \{x + f \colon x \in \mathcal{D}(\Gamma),\ f \in H^2_{\mathcal{U}},\ P_{\mathcal{Z}} f = \Gamma x\} \\
&= \{x + P_{\mathcal{Z}} f + P_{H^2_{\mathcal{U}} \cap \mathcal{M}} f \colon P_{\mathcal{Z}} f = \Gamma x\} \\
&= \{x + \Gamma x \colon x \in \mathcal{D}(\Gamma)\} \oplus (H^2_{\mathcal{U}} \cap \mathcal{M}) \\
&= \mathcal{M}.
\end{aligned}$$

Suppose next that $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ for the admissible Sylvester data set $\mathfrak{S}$ as in (4.7). Then $\mathcal{M}$ being simply invariant for $M_z$ is equivalent to the operator $(\widehat{Z}_{\mathcal{M}})^* \colon \mathcal{M} \to \mathcal{M}$ given by

$$\widehat{Z}^*_{\mathcal{M}} = P_{\mathcal{M}} (M_z)^*|_{\mathcal{M}} \tag{4.12}$$

being strongly stable, i.e., to

$$\|P_{\mathcal{M}} (M_{z^{-1}})^n f\| \to 0 \text{ for each } f \in \mathcal{M}.$$

Since $P_{H^2_{\mathcal{U}}} M_{z^{-1}}|_{H^2_{\mathcal{U}}}$ is strongly stable, an equivalent condition is that the operator

$$P_{L^2\text{-clos}\,(\mathcal{M} + H^2_{\mathcal{U}})} (M_z)^*|_{L^2\text{-clos}\,(\mathcal{M} + H^2_{\mathcal{U}})} \tag{4.13}$$

is strongly stable, where

$$L^2\text{-clos}\,(\mathcal{M} + H^2_{\mathcal{U}}) = \mathcal{M}_{(C,A,0,0,0)} = \operatorname{Ran} \widehat{\mathcal{O}}^b_{C,A} \oplus H^2_{\mathcal{U}}.$$

Let us note that the identity

$$P_{\operatorname{Ran} \widehat{\mathcal{O}}^b_{C,A}} M_z \widehat{\mathcal{O}}^b_{C,A} = P_{H^{2\perp}_{\mathcal{U}}} M_z \widehat{\mathcal{O}}^b_{C,A} = \widehat{\mathcal{O}}^b_{C,A} A$$

implies that the bounded, invertible operator

$$X := \widehat{\mathcal{O}}^b_{C,A} \colon \mathcal{X}_{\mathcal{P}} \to \operatorname{Ran} \widehat{\mathcal{O}}^b_{C,A}$$

implements a similarity between $\widehat{A} := P_{\operatorname{Ran} \widehat{\mathcal{O}}^b_{C,A}} M_z|_{\operatorname{Ran} \widehat{\mathcal{O}}^b_{C,A}}$ and $A$:

$$\widehat{A} X = X A.$$

Then also $\widehat{A}^* = X^{*-1} A^* X^*$. Thus the strong stability of

$$\widehat{A}^* = P_{\operatorname{Ran} \widehat{\mathcal{O}}^b_{C,A}} M_{z^{-1}}|_{\operatorname{Ran} \widehat{\mathcal{O}}^b_{C,A}}$$

is equivalent to strong stability of $A^*$. Thus statement (1a) in Theorem 4.2 follows once we establish the following lemma.

**Lemma 4.3.** *The following are equivalent.*

1. *The operator given by* (4.13) *is strongly stable.*
2. *The operator* $P_{\text{Ran } \widehat{\mathcal{O}}^b_{C,A}} M_{z^{-1}}|_{\text{Ran } \widehat{\mathcal{O}}^b_{C,A}}$ *is strongly stable.*

*Proof of Lemma* 4.3. Condition (1) in Lemma 4.3 means that

$$\lim_{n\to\infty} \left\| P_{\text{Ran } \widehat{\mathcal{O}}^b_{C,A} \oplus H^2_{\mathcal{U}}} (M_{z^{-1}})^n f \right\| = 0 \text{ for all } f \in \text{Ran } \widehat{\mathcal{O}}^b_{C,A} \oplus H^2_{\mathcal{U}}$$

while condition (2) means that

$$\lim_{n\to\infty} \left\| P_{\text{Ran } \widehat{\mathcal{O}}^b_{C,A}} (M_{z^{-1}})^n f \right\| = 0 \text{ for all } f \in \text{Ran } \widehat{\mathcal{O}}^b_{C,A}.$$

Thus (2) follows from (1) simply by restricting $f \in \text{Ran } \widehat{\mathcal{O}}^b_{C,A} \oplus H^2_{\mathcal{U}}$ to $f \in \text{Ran } \widehat{\mathcal{O}}^b_{C,A}$.

For the converse, note that condition (2) means that the $M_{z^{-1}}$-invariant subspace $H^{2\perp}_{\mathcal{U}} \ominus \text{Ran } \widehat{\mathcal{O}}^b_{C,A}$ has a Beurling-Lax representation as

$$H^{2\perp}_{\mathcal{U}} \ominus \text{Ran}\, \widehat{\mathcal{O}}^b_{C,A} = \Theta^* H^{2\perp}_{\mathcal{U}}$$

where $\Theta$ is a *two-sided* inner function in $H^\infty_{\mathcal{L}(\mathcal{U})}$ (i.e., both $\Theta(z)$ and $\Theta(1/\overline{z})^*$ are isometric operators for $|dz|$-a.e. $z \in \mathbb{T}$) (see [38]). Thus

$$\text{Ran } \widehat{\mathcal{O}}^b_{C,A} = H^{2\perp}_{\mathcal{U}} \ominus \Theta^* H^{2\perp}_{\mathcal{U}} = \Theta^* H^2_{\mathcal{U}} \ominus H^2_{\mathcal{U}}$$

and

$$\text{Ran } \widehat{\mathcal{O}}^b_{c,A} \oplus H^2_{\mathcal{U}} = \Theta^* H^2_{\mathcal{U}}.$$

From this representation it is clear that $P_{\text{Ran } \widehat{\mathcal{O}}^b_{C,A} \oplus H^2_{\mathcal{U}}} (M_z)^*|_{\text{Ran } \widehat{\mathcal{O}}^b_{C,A} \oplus H^2_{\mathcal{U}}}$ is strongly stable.

□

We next tackle statement (1b) in Theorem 4.2. Note that the $M_z$-invariant subspace $\mathcal{M}$ is full range if and only if $\mathcal{M}^\perp$ is simply invariant for $M_{z^{-1}}$. By a computation to come (see Proposition 6.2 below), if $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ for the admissible Sylvester data set $\mathfrak{S} = (C, A, , Z, B, \Gamma)$, then $\mathcal{M}^\perp = \mathcal{M}^\times_{\mathfrak{S}^\perp}$ where $\mathfrak{S}^\perp$ is the admissible Sylvester set given by

$$\mathfrak{S}^\perp = (B^*, Z^*, A^*, C^*, -\Gamma^*).$$

In this way we see that statement (1b) in Theorem 4.2 follows directly from statement (2a). This concludes the proof of statement (1) in Theorem 4.2. □

*Proof of (2) in Theorem* 4.2. The proof of these statements is exactly parallel to the proof of the parallel assertions in part (1) of the theorem; hence we leave the details to the reader. □

*Proof of (3) in Theorem* 4.2. Assume that $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ and $\mathcal{M}^\times = \mathcal{M}^\times_{\mathfrak{S}^\times}$ for two admissible Sylvester data sets $\mathfrak{S}$ and $\mathfrak{S}^\times$ as in (4.5). To verify statement (3) in Theorem 4.2, it suffices to verify the general statements (4.10) and (4.11). By the definition of $\mathcal{M}_{\mathfrak{S}}$ in (4.2), we read off that

$$\mathcal{M} \cap H^2_{\mathcal{U}} = \text{Ker } \widehat{\mathcal{C}}^b_{Z,B}.$$

Since by assumption $(Z,B)$ is exactly controllable, we know that $\mathcal{G}_{Z,B}$ is strictly positive definite and that

$$\mathcal{I}_{Z,B} := (\mathcal{G}_{Z,B})^{-1/2}\widehat{\mathcal{C}}^b_{Z,B} \colon H^2_{\mathcal{U}} \to \mathcal{X}_{\mathcal{Z}}$$

is an isometry with kernel equal to $\mathcal{M} \cap H^2_{\mathcal{U}}$. Then the orthogonal projection onto $H^2_{\mathcal{U}} \ominus (H^2_{\mathcal{U}} \cap \mathcal{M}) = (\text{Ker } \mathcal{I}_{Z,B})^\perp$ is given by

$$P_{H^2_{\mathcal{U}} \ominus (H^2_{\mathcal{U}} \cap \mathcal{M})} = (\mathcal{I}_{Z,B})^* \mathcal{I}_{Z,B} = (\widehat{\mathcal{C}}^b_{Z,B})^* (\mathcal{G}_{Z,B})^{-1} \widehat{\mathcal{C}}^b_{Z,B}.$$

In case $\widehat{\mathcal{O}}^b_{C,Z}x + f$ is a generic element of $\mathcal{M}$ (so $x \in \mathcal{D}(\Gamma)$ and $\widehat{\mathcal{C}}^b_{Z,B} f = \Gamma x$), we therefore have

$$P_{H^2_{\mathcal{U}} \ominus (H^2_{\mathcal{U}} \cap \mathcal{M})} f = (\widehat{\mathcal{C}}^b_{Z,B})^* (\mathcal{G}_{Z,B})^{-1} \mathcal{C}_{Z,B} f = (\widehat{\mathcal{C}}^b_{Z,B})^* (\mathcal{G}_{Z,B})^{-1} \Gamma x.$$

We now conclude that

$$\mathcal{M} = \left( \widehat{\mathcal{O}}^b_{C,A} + (\widehat{\mathcal{C}}^b_{Z,B})^* \mathcal{G}^{-1}_{Z,B} \Gamma \right) \mathcal{D}(\Gamma) \oplus (M \cap H^2_{\mathcal{U}}). \tag{4.14}$$

By an exactly parallel analysis one can show that

$$\mathcal{M}^\times = (\mathcal{M}^\times \cap H^{2\perp}_{\mathcal{U}}) \oplus \left( (\widehat{\mathcal{C}}^f_{Z^\times, B^\times})^* (\mathcal{G}_{Z^\times, B^\times})^{-1} \Gamma^\times + \widehat{\mathcal{O}}^f_{C^\times, A^\times} \right) \mathcal{D}(\Gamma^\times). \tag{4.15}$$

Note also that the operator $\widehat{\mathcal{O}}^b_{C,A} + (\widehat{\mathcal{C}}^b_{Z,B})^* \mathcal{G}^{-1}_{Z,B}\Gamma$ is injective on $\mathcal{D}(\Gamma)$ since $(C,A)$ is observable and Ran $\widehat{\mathcal{O}}^b_{C,A} \subset H^{2\perp}_{\mathcal{U}}$ while Ran $(\widehat{\mathcal{C}}^b_{Z,B})^* \subset H^2_{\mathcal{U}}$, and hence (4.14) sets up a one-to-one correspondence between $\mathcal{D}(\Gamma) \times (H^2_{\mathcal{U}} \cap \mathcal{M})$ and $\mathcal{M}$. Similarly, (4.15) sets up a one-to-one correspondence between $\mathcal{D}(\Gamma^\times) \times (H^{2\perp}_{\mathcal{U}} \cap \mathcal{M}^\times)$ and $\mathcal{M}^\times$.

Define an operator $\widehat{\Gamma}_{\mathfrak{S},\mathfrak{S}^\times} \colon \mathcal{D}(\Gamma) \oplus \mathcal{D}(\Gamma^\times) \to L^2_{\mathcal{U}}$ by

$$\widehat{\Gamma}_{\mathfrak{S},\mathfrak{S}^\times} = \begin{bmatrix} \widehat{\mathcal{O}}^b_{C,A} + (\widehat{\mathcal{C}}^b_{Z,B})^* (\mathcal{G}_{Z,B})^{-1}\Gamma & (\widehat{\mathcal{C}}^f_{Z^\times,B^\times})^* (\mathcal{G}_{Z^\times,B^\times})^{-1}\Gamma^\times + \widehat{\mathcal{O}}^f_{C^\times,A^\times} \end{bmatrix}. \tag{4.16}$$

Note that, by (4.14) and (4.15), we have

$$\mathcal{M} + \mathcal{M}^\times = \text{Ran } \widehat{\Gamma}_{\mathfrak{S},\mathfrak{S}^\times} + (H^{2\perp}_{\mathcal{U}} \cap \mathcal{M}^\times) + (H^2_{\mathcal{U}} \cap \mathcal{M}). \tag{4.17}$$

If we extend $\widehat{\mathcal{C}}^f_{Z^\times,B^\times} \colon H^{2\perp}_{\mathcal{U}} \to \mathcal{X}^\times_{\mathcal{Z}}$ and $\widehat{\mathcal{C}}^b_{Z,B} \colon H^2_{\mathcal{U}} \to \mathcal{X}_{\mathcal{Z}}$ to act linearly on all of $L^2_{\mathcal{U}}$ by setting

$$\widehat{\mathcal{C}}^f_{Z^\times,B^\times}|_{H^2_{\mathcal{U}}} = 0, \qquad \widehat{\mathcal{C}}^b_{Z,B}|_{H^{2\perp}_{\mathcal{U}}} = 0$$

then we have the factorization

$$\Gamma_{\mathfrak{S},\mathfrak{S}^\times} = \begin{bmatrix} -\widehat{\mathcal{C}}^f_{Z^\times,B^\times} \\ \widehat{\mathcal{C}}^b_{Z,B} \end{bmatrix} \widehat{\Gamma}_{\mathfrak{S},\mathfrak{S}^\times}. \tag{4.18}$$

Moreover, since

$$\text{Ker}\begin{bmatrix} -\widehat{\mathcal{C}}^f_{Z^\times,B^\times} \\ \widehat{\mathcal{C}}^b_{Z,B} \end{bmatrix} = (H^{2\perp}_{\mathcal{U}} \cap \mathcal{M}^\times) \oplus (H^2_{\mathcal{U}} \cap \mathcal{M}), \tag{4.19}$$

it follows that

$$\begin{aligned}(\mathcal{X}_{\mathcal{Z}} \oplus \mathcal{X}^\times_{\mathcal{Z}})/\,\text{Ran}\,\Gamma_{\mathfrak{S},\mathfrak{S}^\times} &\cong L^2_{\mathcal{U}}/\left(\text{Ran}\,\widehat{\Gamma}_{\mathfrak{S},\mathfrak{S}^\times} + (H^{2\perp}_{\mathcal{U}} \cap \mathcal{M}^\times) + (H^2 \cap \mathcal{M})\right) \\ &= L^2_{\mathcal{U}}/(\mathcal{M}^\times + \mathcal{M}) \end{aligned} \tag{4.20}$$

and (4.11) is established.

Suppose now that $f \in \mathcal{M} \cap \mathcal{M}^\times$. Write $f = g + h$ with $g \in H^{2\perp}_{\mathcal{U}}$ and $h \in H^2_{\mathcal{U}}$. Since $f \in \mathcal{M}$, necessarily

$$g = \widehat{\mathcal{O}}^b_{C,A}x \text{ for some } x \in \mathcal{D}(\Gamma) \text{ and } \widehat{\mathcal{C}}^b_{Z,B}h = \Gamma x. \tag{4.21}$$

Since also $f \in \mathcal{M}^\perp$, necessarily

$$h = \widehat{\mathcal{O}}^f_{C^\times,A^\times}y \text{ for some } y \in \mathcal{D}(\Gamma^\times) \text{ and } \widehat{\mathcal{C}}^f_{Z^\times,B^\times}g = \Gamma^\times y. \tag{4.22}$$

Using the first parts of (4.21) and (4.22) to eliminate $g$ and $h$ leaves us with

$$f = \widehat{\mathcal{O}}^b_{C,A}x + \widehat{\mathcal{O}}^f_{C^\times,A^\times}y \text{ where } \begin{cases} \widehat{\mathcal{C}}^b_{Z,B}\widehat{\mathcal{O}}^f_{C^\times,A^\times}y = \Gamma x \\ \widehat{\mathcal{C}}^f_{Z^\times,B^\times}\widehat{\mathcal{O}}^b_{C,A}y = \Gamma^\times y, \end{cases}$$

i.e.,

$$f = \begin{bmatrix} \widehat{\mathcal{O}}^b_{C,A} & -\mathcal{O}^f_{C^\times,A^\times} \end{bmatrix} \begin{bmatrix} x \\ -y \end{bmatrix} \text{ where } \begin{bmatrix} x \\ -y \end{bmatrix} \in \text{Ker}\,\Gamma_{\mathfrak{S},\mathfrak{S}^\times}. \tag{4.23}$$

Conversely, it follows from the definitions of $\mathcal{M}$ and $\mathcal{M}^\times$ that any $f$ of the form (4.23) is in $\mathcal{M} \bigcap \mathcal{M}^\times$. We have established: *the operator*

$$\begin{bmatrix} \widehat{\mathcal{O}}^b_{C,A} & -\widehat{\mathcal{O}}^f_{C^\times,A^\times} \end{bmatrix}\Big|_{\text{Ker}\,\Gamma_{\mathfrak{S},\mathfrak{S}^\times}} : \text{Ker}\,\Gamma_{\mathfrak{S},\mathfrak{S}^\times} \to \mathcal{M} \cap \mathcal{M}^\times$$

*establishes a bijective correspondence between* Ker $\Gamma_{\mathfrak{S},\mathfrak{S}^\times}$ *and* $\mathcal{M} \cap \mathcal{M}^\times$. This establishes (4.10) and completes the proof of Theorem 4.2 □

If we assume that dim $\mathcal{U}$ is finite, we can use the construction in the proof of Lemma 4.3 to say more about $L^2$-regular $\mathcal{L}(\mathcal{U})$-valued functions. As preparation for the statement of the result, we recall the following definitions (see [28, 29]). Given a measurable $\mathcal{L}(\mathcal{U})$-valued function defined on $\mathbb{T}$, we shall say that $W$ *has meromorphic pseudocontinuation to the unit disk* $\mathbb{D}$ if there is a meromorphic $\mathcal{L}(\mathcal{U})$-valued function $\widetilde{W}$ on $\mathbb{D}$ so that $W(z) = \lim_{\lambda\to z}\widetilde{W}(\lambda)$ for a.e. $z \in \mathbb{T}$ where the limit is taken as $\lambda \in \mathbb{D}$ approaches $z \in \mathbb{T}$ nontangentially. We say that $W$ *has pseudocontinuation of bounded type to* $\mathbb{D}$ if one can take the function $\widetilde{W}(\lambda)$ to have the form $W(\lambda) = H(\lambda)/\phi(\lambda)$ where $H(\lambda)$ is a bounded holomorphic $\mathcal{L}(\mathcal{U})$-valued function and $\phi(z)$ is a nonzero bounded holomorphic scalar function on $\mathbb{D}$. Similarly, we say that $W$ *has meromorphic pseudocontinuation to* $\mathbb{D}_e := (\mathbb{C} \cup \{\infty\}) \setminus \overline{\mathbb{D}}$ or $W$ *has pseudocontinuation of bounded type to* $\mathbb{D}_e$ if one replaces the unit disk $\mathbb{D}$ by the exterior of the unit disk $\mathbb{D}_e$ in the preceding definitions.

**Theorem 4.4.** *Suppose that $W \in L^2_{\mathcal{L}(\mathcal{U})}$ is $L^2$-regular. Then there are two-sided inner functions $\Theta, \Phi$ such that*

$$\Theta W \in H^2_{\mathcal{L}(\mathcal{U})}, \qquad \Phi^{-1} W \in H^{2\perp}_{\mathcal{L}(\mathcal{U})}. \tag{4.24}$$

*In particular, if $\dim \mathcal{U} < \infty$, then $W$ has pseudocontinuation of bounded type to both $\mathbb{D}$ and to $\mathbb{D}_e$.*

*Proof.* Suppose that $W$ is $L^2$-regular. Then by Theorem 3.4 we know that $\mathcal{M} = L^2\text{-clos}\, W H^\infty_{\mathcal{U}}$ is full-range simply-invariant for $M_z$. As explained in the proof of Lemma 4.3, there is a two-sided inner function $\Theta$ so that $\Theta\mathcal{M} \subset H^2_{\mathcal{U}}$. As $\mathcal{M} = L^2\text{-clos}\, W H^\infty_{\mathcal{U}}$, it then follows that $\Theta \cdot W \in H^2_{\mathcal{L}(\mathcal{U})}$.

Now suppose that $\dim \mathcal{U} < \infty$. Let us denote by $G$ the $H^2_{\mathcal{L}(\mathcal{U})}$-function given by

$$G = \Theta \cdot W.$$

Solving for $W$ then gives

$$W(z) = \Theta(z)^{-1} G(z) = (\operatorname{adj} \Theta(z)) G(z) / \det(\Theta(z)). \tag{4.25}$$

Here $\operatorname{adj} \Theta(z)$ is the adjugate matrix for $\Theta(z)$ appearing in Cramer's rule for the inverse. Since $\Theta(z)$ is bounded and holomorphic on $\mathbb{D}$, so also is $\operatorname{adj} \Theta(z)$ as well as $\det \Theta(z)$. It is known that $H^2$-functions are of bounded type on the unit disk (i.e., can be written as the ratio of bounded holomorphic functions). Hence (4.25) gives the pseudocontinuation of $W(z)$ to $\mathbb{D}$.

To verify the statement concerning $\Phi$, we use Theorem 3.4 to see that $\mathcal{M}^\times = L^2\text{-clos}\, W \overline{H^\infty_{\mathcal{U},0}}$ is full-range and simply-invariant for $M_{z^{-1}}$. Hence the argument in Lemma 4.3 can be used to show that there is a two-sided inner function $\Psi$ so that $\Psi^{-1}\mathcal{M}^\times \subset H^{2\perp}_{\mathcal{U}}$. This then implies that $\Psi^{-1} W =: H \in H^{2\perp}_{\mathcal{L}(\mathcal{U})}$. One can now proceed as before (with $\mathbb{D}_e$ playing the role of $\mathbb{D}$) to see that $W$ has pseudocontinuation of bounded type to $\mathbb{D}_e$ for the case where $\dim \mathcal{U} < \infty$. □

If $W$ has a strongly bi-dichotomous realization as in Definition 2.2, then in particular both $W$ and $W^{-1}$ are in $L^2_{\mathcal{L}(\mathcal{U})}$. Hence, in case $W$ also satisfies the regularity condition in part (2) of Definition 3.3, then $W$ is $L^2$-regular and Theorem 3.4 applies. Our next result identifies how to compute the pole-zero data $\mathfrak{S}, \mathfrak{S}^\times$ for the associated dual shift-invariant pair of subspaces $(\mathcal{M}, \mathcal{M}^\times)$ directly from the realization data.

**Theorem 4.5.** *Suppose that $W \in L^2_{\mathcal{L}(\mathcal{U})}$ has a strongly bi-dichotomous realization as in Definition 2.2, so (2.18), (2.19) holds with the realization data*

$$\mathfrak{D} = (C_\pm, A_\pm, B_\pm, D, C^\times_\pm, A^\times_\pm, B^\times_\pm, D^\times)$$

*satisfying conditions (1) -(5) in Definition 2.2. Let $\mathcal{M}$ and $\mathcal{M}^\times$ be the associated forward- and backward-shift-invariant subspaces*

$$\mathcal{M} = L^2\text{-}clos\ W \cdot H^\infty_{\mathcal{U}}, \qquad \mathcal{M}^\times = L^2\text{-}clos\ W \cdot \overline{H^\infty_{\mathcal{U},0}}.$$

*Then there exist coupling operators*

$$\Gamma\colon \mathcal{D}(\Gamma) \subset \mathcal{X}_+ \to \mathcal{X}_+^\times, \qquad \Gamma^\times\colon \mathcal{D}(\Gamma^\times) \subset \mathcal{X}_- \to \mathcal{X}_-^\times$$

*so that*

$$\mathfrak{S} = (C_-, A_-, A_-^\times, B_-^\times, \Gamma), \qquad \mathfrak{S}^\times = (C_+, A_+, A_+^\times, B_+^\times, \Gamma^\times) \tag{4.26}$$

*are admissible Sylvester data sets such that*

$$\mathcal{M} = \mathcal{M}_{\mathfrak{S}}, \qquad \mathcal{M}^\times = \mathcal{M}^\times_{\mathfrak{S}^\times}$$

*(see* (4.3) *and* (4.4)*).*

*Remark* 4.6. Note that the cross-decoration $\times$ for the matrices in a bi-dichotomous realization refers to matrices associated with the realization for $W^{-1}$ whereas those without the $\times$ are matrices associated with the realization for $W$. In connection with Sylvester data sets, on the other hand, matrices with the cross-decoration $\times$ refer to matrices associated with the backward-shift-invariant subspace $\mathcal{M}^\times$ while those without the decoration $\times$ refer to data for the forward-shift-invariant subspace $\mathcal{M}$. Theorem 4.5 is one place where these conventions get tangled up; this explains the perhaps startling appearance of $\mathfrak{S}$ and $\mathfrak{S}^\times$ in (4.26). A similar tangling occurs in the context of Theorem 5.1. Note that matrices $A, B, C$ coming from a bi-dichotomous realization always carry a subscript $A_\pm, B_\pm, C_\pm$ (or $A_\pm^\times, B_\pm^\times, C_\pm^\times$), with $+$ associated with the part analytic on the unit disk $\mathbb{D}$ and $-$ associated with the part analytic on the exterior of the closed unit disk, while $C, A, Z, B$ from a Sylvester data set carry no subscripts.

*Proof of Theorem* 4.5. We verify only the statement concerning $\mathcal{M}$ and $\mathfrak{S}$ as the proof of the corresponding statements concerning $\mathcal{M}^\times$ and $\mathfrak{S}^\times$ is exactly parallel.

We first verify that

$$P_{H^{2\perp}_{\mathcal{U}}} \mathcal{M} = \{C_-(zI - A_-)^{-1}x \colon x \in \mathcal{D}\} \tag{4.27}$$

for a dense linear submanifold $\mathcal{D}$ of $\mathcal{X}_-$. Consider first the case where $h = W \cdot p \in \mathcal{M}$ for a polynomial $p(z) = \sum_{n=0}^\infty p_n z^n$ where $p_n \in \mathcal{U}$ is zero once $n \geq N$ for some $N < \infty$. Then we compute, for $n < 0$,

$$\begin{aligned} [W \cdot p]_n &= \sum_{i,j \in \mathbb{Z}_+ \colon -i-1+j=n} C_- A_-^i B_- p_j \\ &= \sum_{j \in \mathbb{Z}_+} C_- A_-^{-n-1+j} B_- p_j \\ &= C_- A_-^{-n-1} \cdot \sum_{j \in \mathbb{Z}_+} A_-^j B_- p_j \\ &= C_- A_-^{-n-1} x \text{ where } x = \sum_{j \in \mathbb{Z}_+} A_-^j B_- p_j = \widehat{\mathcal{C}}^b_{A_-, B_-} p. \end{aligned}$$

(Note that the sum defining $x$ is finite since $p(z)$ is a polynomial.) We conclude that

$$P_{H^{2\perp}_{\mathcal{U}}} (W(z) \cdot p(z)) = C_-(zI - A_-)^{-1} x = \widehat{\mathcal{O}}^b_{C_-, A_-} x$$

where $x \in \operatorname{Ran} \widehat{\mathcal{C}}_{A_-,B_-}|_{\mathcal{P}_\mathcal{U}}$. Conversely, by reversing the above computation we see that

$$\left\{\widehat{\mathcal{O}}^b_{C_-,A_-} x \colon x \in \operatorname{Ran} \widehat{\mathcal{C}}_{A_-,B_-}|_{\mathcal{P}_\mathcal{U}}\right\} \subset P_{H^{2\perp}_\mathcal{U}} \mathcal{M}. \tag{4.28}$$

Now suppose that $h \in L^2_\mathcal{U}$ is a general element of $\mathcal{M}$. Then $h = L^2_\mathcal{U} - \lim_{n\to\infty} W \cdot p_n$ for a sequence of polynomials $p_n \in \mathcal{P}_\mathcal{U}$. Then also

$$P_{H^{2\perp}_\mathcal{U}} h = L^2_\mathcal{U} - \lim_{n\to\infty} P_{H^{2\perp}_\mathcal{U}}(W \cdot p_n).$$

By the polynomial case done above, we know that

$$P_{H^{2\perp}_\mathcal{U}}(W \cdot p_n) = \widehat{\mathcal{O}}_{C_-,A_-} x_n$$

for a vector $x_n \in \mathcal{X}_-$. Since $P_{H^{2\perp}_\mathcal{U}}(W \cdot p_n) \to P_{H^{2\perp}_\mathcal{U}} h$ in $H^{2\perp}_\mathcal{U}$ as $n \to \infty$, by the assumed exact observability of $(C_-, A_-)$ it follows that $\lim_{n\to\infty} x_n = x$ exists in $\mathcal{X}$ and then also

$$P_{H^{2\perp}_\mathcal{U}} h = \widehat{\mathcal{O}}^b_{C_-,A_-} x.$$

Thus there is a linear submanifold $\mathcal{D}$ of $\mathcal{X}_-$ so that (4.27) holds. As $\mathcal{D}$ contains $\operatorname{Ran} \widehat{\mathcal{C}}^b_{A_-,B_-}|_{\mathcal{P}_\mathcal{U}}$ by (4.28) and $(A_-, B_-)$ is assumed to be controllable, it follows that $\mathcal{D}$ is dense in $\mathcal{X}_-$.

We next verify that

$$\mathcal{M} \cap H^2_\mathcal{U} = \operatorname{Ker} \widehat{\mathcal{C}}_{A^\times_-,B^\times_-}. \tag{4.29}$$

Suppose therefore that $h \in \mathcal{M} \cap H^2_\mathcal{U}$. Then

$$h = L^2_\mathcal{U} - \lim_{n\to\infty} W \cdot p_n$$

for a sequence of analytic polynomials $p_n \in \mathcal{P}_\mathcal{U}$. Since $W^{-1} \in L^2_{\mathcal{L}(\mathcal{U})}$, $M_{W^{-1}} \colon f \mapsto W^{-1} \cdot f$ is continuous from $L^2_\mathcal{U}$ into $L^1_\mathcal{U}$ in the weak sense of Proposition 2.1, and hence

$$W^{-1} \cdot h = H^1_\mathcal{U} - \lim_{n\to\infty} W^{-1} \cdot (W p_n) = H^1_\mathcal{U} - \lim_{n\to\infty} p_n$$

exists as a weak limit in $H^1_\mathcal{U}$. In particular, $W^{-1} h \in H^1_\mathcal{U}$ and it follows that

$$[W^{-1} \cdot h]_n = 0 \text{ for } n \in \mathbb{Z}_-.$$

We compute, for $n \in \mathbb{Z}_-$,

$$\begin{aligned} 0 = [W^{-1} \cdot h]_n &= [C^\times_-(zI - A^\times_-)^{-1} B^\times_- h(z)]_n \\ &= \sum_{i,j\in\mathbb{Z}_+ \colon -i-1+j=n} C^\times_-(A^\times_-)^i B^\times_- h_j \\ &= \sum_{j\in\mathbb{Z}_+} C^\times_- A^{\times(-n-1+j)} B^\times_- h_j \\ &= C^\times_- A^{\times(-n-1)} \cdot \sum_{j\in\mathbb{Z}_+} (A^\times_-)^j B^\times_- h_j \\ &= C^\times_-(A^\times)^{-n-1} \cdot \widehat{\mathcal{C}}^b_{A^\times_-,B^\times_-} h. \end{aligned}$$

Since by assumption $(C_-^\times, A_-^\times)$ is observable, we conclude that $\widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times} h = 0$. Conversely, reversing the argument shows that if $h \in \operatorname{Ker} \widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times}$, then $W^{-1} \cdot h \in H^1_{\mathcal{U}}$ and hence $h \in H^2_{\mathcal{U}} \cap \mathcal{M}$. In this way we have verified (4.29).

Now let $k$ be an arbitrary element of $\mathcal{M}$. By (4.27) $k$ has the form

$$k = \widehat{\mathcal{O}}^b_{C_-, A_-} x + f$$

for some $x \in \mathcal{D}$ and $f \in H^2_{\mathcal{U}}$. Define an operator $\Gamma \colon \mathcal{D} \to \mathcal{X}_-^\times$ by

$$\Gamma x = \widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times} f. \tag{4.30}$$

Since by (4.29) we know that $\mathcal{M} \cap H^2_{\mathcal{U}}$ is exactly equal to Ker $\widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times}$, it follows that the formula (4.30) for $\Gamma$ is well-defined, i.e., the expression $\widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times} f$ is independent of the choice of $f \in H^2_{\mathcal{U}}$ for which $\widehat{\mathcal{O}}^b_{C_-, A_-} x + f \in \mathcal{M}$. It follows that we recover $\mathcal{M}$ from $\mathfrak{S} = (C_-, A_-, A_-^\times, B_-^\times, \Gamma)$ (as in (4.26)) as $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$.

To check that $\mathfrak{S}$ given by (4.26) is an admissible Sylvester data set, it remains only to check that $\Gamma$ with domain $\mathcal{D}$ as given above is a closed operator. This follows from $\mathcal{M}$ being a closed subspace of $L^2_{\mathcal{U}}$ as follows (the reverse direction was done in the proof of part (1) of Theorem 4.2). Suppose that $\{x_n\}_{n \in \mathbb{Z}_+}$ is a sequence of vectors from $\mathcal{D}$ such that

$$\lim_{n \to \infty} x_n = x \in \mathcal{X}_+ \text{ and } \lim_{n \to \infty} \Gamma x_n = y \in \mathcal{X}_+^\times.$$

As $\widehat{\mathcal{C}}^b_{C_-, A_-}$ is bounded, we then get

$$\lim_{n \to \infty} \widehat{\mathcal{C}}^b_{C_-, A_-} x_n = \widehat{\mathcal{C}}^b_{C_-, A_-} x. \tag{4.31}$$

Since $(A_-^\times, B_-^\times)$ is exactly controllable, we may choose

$$f_n \in \operatorname{Ran} (\widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times})^* = (\operatorname{Ker} \widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times})^\perp \subset H^2_{\mathcal{U}}$$

so that

$$\Gamma x_n = \widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times} f_n.$$

The characterization of $\mathcal{M}$ as $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ then implies that

$$\widehat{\mathcal{O}}^b_{C_-, A_-} x_n + f_n \in \mathcal{M} \text{ for all } n \in \mathbb{Z}_+. \tag{4.32}$$

The assumption that $\lim_{n \to \infty} \Gamma x_n = y$ exists combined with the exact controllability of $(A_-^\times, B_-^\times)$ implies that

$$\lim_{n \to \infty} f_n = f \in H^2_{\mathcal{U}} \text{ exists in } H^2_{\mathcal{U}} \text{ and } y = \lim_{n \to \infty} \Gamma x_n = \widehat{\mathcal{C}}^b_{A_-^\times, B_-^\times} f. \tag{4.33}$$

But then, by (4.31) and (4.33) we have

$$\lim_{n \to \infty} \left[ \widehat{\mathcal{C}}^b_{C_-, A_-} x_n + f_n \right] = \widehat{\mathcal{C}}^b_{C_-, A_-} x + f. \tag{4.34}$$

By (4.32) and the hypothesis that $\mathcal{M}$ is closed, we conclude that

$$\widehat{\mathcal{C}}^b_{C_-,A_-}x + f \in \mathcal{M}.$$

Now the characterization of $\mathcal{M}$ as $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ together with (4.33) gives us that

$$x \in \mathcal{D} \text{ and } y = \widehat{\mathcal{C}}^b_{A_-^\times,B_-^\times} f = \Gamma x.$$

We have now verified that $\Gamma$ is closed as wanted. This completes the proof of Theorem 4.5. □

## 5. Dichotomous realizations for strongly regular matrix-valued functions via pole-zero data

The following result provides a state-space implementation of Theorem 3.4 as well as a converse to Theorem 4.5. We refer the reader to Remark 4.6 for a guide through the thicket of notation in the following theorem.

**Theorem 5.1.** *Suppose that $(\mathcal{M}, \mathcal{M}^\times)$ is a dual shift-invariant pair of subspaces as in Definition 3.1. Then:*

1. *There are admissible Sylvester data sets*

$$\mathfrak{S} = (C, A, Z, B, \Gamma), \qquad \mathfrak{S}^\times = (C^\times, A^\times, Z^\times, B^\times, \Gamma^\times) \tag{5.1}$$

*such that*

(a) *$A^*$, $Z$, $(A^\times)^*$, $Z^\times$ are all strongly stable (and hence in fact $A$, $Z$, $A^\times$ and $Z^\times$ are all strongly bi-stable),*

(b) *the coupling operator $\Gamma_{\mathfrak{S},\mathfrak{S}^\times}$ (see (4.6)) is invertible, and*

(c) *$\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$, $\mathcal{M}^\times = \mathcal{M}^\times_{\mathfrak{S}^\times}$ (see (4.3) and (4.4)).*

2. *There is an $L^2$-regular $\mathcal{L}(\mathcal{U})$-valued function $W$ on $\mathbb{T}$ such that*

$$\mathcal{M} = L^2\text{-}clos\ W H^\infty_{\mathcal{U}}, \qquad \mathcal{M}^\times = L^2\text{-}clos\ W \overline{H^\infty_{\mathcal{U},0}}. \tag{5.2}$$

*Realization formulas for $W$ and $W^{-1}$ can be constructed from the Sylvester data sets $\mathfrak{S}$ and $\mathfrak{S}^\times$ in (5.1) as follows. Define operators $\Psi$ and $\widetilde{\Psi}$ according to*

$$\Psi = \begin{bmatrix} \Gamma & -Z\widehat{\mathcal{C}}^b_{Z,B}\widehat{\mathcal{O}}^f_{C^\times,A^\times} & -B \\ -Z^\times\widehat{\mathcal{C}}^f_{Z^\times,B^\times}\widehat{\mathcal{O}}^b_{C,A} & \Gamma^\times & -B^\times \end{bmatrix} : \begin{bmatrix} \mathcal{D}(\Gamma) \\ \mathcal{D}(\Gamma^\times) \\ \mathcal{U} \end{bmatrix} \to \begin{bmatrix} \mathcal{X}_{\mathcal{Z}} \\ \mathcal{X}^\times_{\mathcal{Z}} \end{bmatrix} \tag{5.3}$$

$$\widetilde{\Psi} = \begin{bmatrix} -\Gamma^\times & -\widehat{\mathcal{C}}^f_{Z^\times,B^\times}\widehat{\mathcal{O}}^b_{C,A}A \\ -\widehat{\mathcal{C}}^b_{Z,B}\widehat{\mathcal{O}}^f_{C^\times,A^\times}A^\times & -\Gamma \\ -C^\times & -C \end{bmatrix} : \begin{bmatrix} \mathcal{D}(\Gamma^\times) \\ \mathcal{D}(\Gamma) \end{bmatrix} \to \begin{bmatrix} \mathcal{X}^\times_{\mathcal{Z}} \\ \mathcal{X}_{\mathcal{Z}} \\ \mathcal{U} \end{bmatrix}. \tag{5.4}$$

*Then it holds that*

$$\dim \operatorname{Ker} \Psi = \dim \mathcal{U}, \qquad \dim \begin{bmatrix} \mathcal{X}_{\mathcal{Z}} \\ \mathcal{X}^\times_{\mathcal{Z}} \end{bmatrix} / \operatorname{Ran} \widetilde{\Psi} = \dim \mathcal{U}.$$

*Let*

$$\begin{bmatrix} B_- \\ B_+ \\ D \end{bmatrix} : \mathcal{U} \to \text{Ker } \Psi \subset \begin{bmatrix} \mathcal{D}(\Gamma) \\ \mathcal{D}(\Gamma^\times) \\ \mathcal{U} \end{bmatrix} \tag{5.5}$$

*be an isomorphism between* $\mathcal{U}$ *and* Ker $\Psi$*, let*

$$\begin{bmatrix} C_+^\times & C_-^\times & D^\times \end{bmatrix} : \begin{bmatrix} \mathcal{X}_{\mathcal{Z}}^\times \\ \mathcal{X}_{\mathcal{Z}} \\ \mathcal{U} \end{bmatrix} \to \mathcal{U} \tag{5.6}$$

*be a surjective linear map with* Ker $\begin{bmatrix} C_-^\times & C_+^\times & D^\times \end{bmatrix} = \text{Ran } \widetilde{\Psi}$*, and set*

$$L = C_+^\times \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} B_- + C_-^\times \widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{C^\times, A^\times} B_+ + D^\times D. \tag{5.7}$$

*Then $L$ is invertible and one choice of an $L^2$-regular $W$ satisfying* (5.2) *is*

$$W(z) = C(zI - A)^{-1} B_- + D + zC^\times (I - zA^\times)^{-1} B_+ \tag{5.8}$$

*with inverse $W^{-1}$ given by*

$$W(z)^{-1} = L^{-1} \left[ C_-^\times (zI - Z)^{-1} B + D^\times + zC_+^\times (I - zZ^\times)^{-1} B^\times \right] \tag{5.9}$$

*and* (5.8), (5.9) *provide a strongly bi-dichotomous realization for $W$.*

*Proof of (1):* Let $(\mathcal{M}, \mathcal{M}^\times)$ be a dual shift-invariant pair of subspaces of $L^2_{\mathcal{U}}$. In particular, $\mathcal{M}$ is full-range and simply invariant for $M_z$. Hence, by part (1) of Theorem 4.2, $\mathcal{M}$ has the form $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ for an admissible Sylvester data set $\mathfrak{S} = (C, A, Z, B, \Gamma)$ with both $A$ and $Z$ bi-stable. Similarly, as $\mathcal{M}^\times$ is full-range and simply invariant for $M_{z^{-1}}$, part (2) of Theorem 4.2 guarantees that $\mathcal{M}^\times$ has the form $\mathcal{M}^\times = \mathcal{M}^\times_{\mathfrak{S}^\times}$ for an admissible Sylvester data set $\mathfrak{S}^\times = (C^\times, A^\times, Z^\times, B^\times, \Gamma^\times)$ with both $A^\times$ and $Z^\times$ bi-stable. Finally the matching condition (3.1) in part (3) of Definition 3.1 combined with part (3) of Theorem 4.2 guarantees that the coupling operator $\Gamma_{\mathfrak{S}, \mathfrak{S}^\times}$ associated with $\mathfrak{S}$ and $\mathfrak{S}^\times$ as in (4.6) is invertible. We have now verified all of (a), (b), (c) in part (1) of Theorem 5.1. □

*Proof of (2):* We assume that we are given a dual shift-invariant pair of subspaces of the form $\mathcal{M}_{\mathfrak{S}}, \mathcal{M}_{\mathfrak{S}^\times}$ for admissible Sylvester data sets $\mathfrak{S}$, $\mathfrak{S}^\times$ with the stability properties listed in part (1a) in the statement of the theorem. By the proof of Theorem 3.4, we need only construct a function $W \in L^2_{\mathcal{L}(\mathcal{U})}$ so that

$$W\mathcal{U} = z\mathcal{M}^\times \cap \mathcal{M} =: \mathcal{L}.$$

Suppose that $h \in z\mathcal{M}^\times \cap \mathcal{M}$. Then, on the one hand, $h = zk^\times$ where $k^\times$, as an element of $\mathcal{M}^\times$, has the form

$$k^\times(z) = [\mathcal{O}^f_{C^\times, A^\times} x^\times + f^\times](z) = C^\times (I - zA^\times)^{-1} x^\times + f^\times(z)$$

for an $x^\times \in \mathcal{D}(\Gamma^\times)$ and an $f^\times \in H^{2\perp}_{\mathcal{U}}$ subject to

$$\widehat{\mathcal{C}}^f_{Z^\times, B^\times} f^\times = \Gamma^\times x.$$

Then $h = zk^\times$ has the form

$$\begin{aligned} h(z) &= [M_z \cdot (\mathcal{O}^f_{C^\times, A^\times} x^\times + f^\times)](z) \\ &= zC^\times (I - zA^\times)^{-1} x^\times + zf^\times(z) \\ &= zC^\times (I - zA^\times)^{-1} x^\times + (P_{H^{2\perp}_{\mathcal{U}}} M_z f^\times)(z) + [f^\times]_{-1}. \end{aligned} \tag{5.10}$$

On the other hand, as an element of $\mathcal{M}$, $h$ also has the form

$$h(z) = [\widehat{\mathcal{O}}^b_{C,A} x + f](z) = C(zI - A)^{-1} x + f(z) \tag{5.11}$$

where $x \in \mathcal{D}(\Gamma)$ and $f \in H^2_{\mathcal{U}}$ satisfy

$$\widehat{\mathcal{C}}^b_{Z,B} f = \Gamma x.$$

By equating $H^{2\perp}_{\mathcal{U}}$- and $H^2_{\mathcal{U}}$-components in (5.10) and (5.11) we see that

$$\begin{aligned} &[P_{H^{2\perp}_{\mathcal{U}}} M_z f^\times](z) = C(zI - A)^{-1} x, \\ &f(z) = [f^\times]_{-1} + zC^\times (I - zA^\times)^{-1} x^\times. \end{aligned}$$

By combining (5.10) and (5.11) we see that we can assume that $h$ has the form

$$h(z) = C(zI - A)^{-1} x + u + zC^\times (I - zA^\times)^{-1} x^\times \tag{5.12}$$

for an $x \in \mathcal{D}(\Gamma)$, a $u \in \mathcal{U}$ (considered as a constant-function element of $H^2_{\mathcal{U}}$) and an $x^\times \in \mathcal{D}(\Gamma^\times)$. The added condition for an element of the form (5.12) to be in $\mathcal{M}$ is that

$$\Gamma x = \widehat{\mathcal{C}}^b_{Z,B} (u + M_z \widehat{\mathcal{O}}^f_{C^\times, A^\times} x^\times). \tag{5.13}$$

We note that

$$\widehat{\mathcal{C}}^b_{Z,B} u = Bu$$

and we compute

$$\begin{aligned} \widehat{\mathcal{C}}^b_{Z,B} M_z \widehat{\mathcal{O}}^f_{C^\times, A^\times} x^\times &= \sum_{n=0}^\infty Z^n B [M_z \widehat{\mathcal{O}}^f_{C^\times, A^\times} x^\times]_n \\ &= \sum_{n=1}^\infty Z^n B C^\times (A^\times)^{n-1} x^\times \\ &= Z \cdot \sum_{n=0}^\infty Z^n B C^\times (A^\times)^n x^\times \\ &= Z \widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{C^\times, A^\times} x^\times. \end{aligned}$$

Thus (5.13) can be written in more succinct form as

$$\begin{bmatrix} \Gamma & -Z \widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{C^\times, A^\times} & -B \end{bmatrix} \begin{bmatrix} x \\ x^\times \\ u \end{bmatrix} = 0. \tag{5.14}$$

Similarly, we require that $h$ as in (5.12) have the property that

$$z^{-1} \cdot h(z) = z^{-1} C(zI - A)^{-1} x + z^{-1} u + C^\times (I - zA^\times)^{-1} x^\times$$

be in $\mathcal{M}^\times$. The additional constraint for this to happen is that

$$\widehat{\mathcal{C}}^f_{Z^\times,B^\times}(M_{z^{-1}}\widehat{\mathcal{O}}^b_{C,A}x + M_{z^{-1}}u) = \Gamma^\times x^\times. \tag{5.15}$$

We note that

$$\widehat{\mathcal{C}}^f_{Z^\times,B^\times}M_{z^{-1}}u = B^\times u$$

and a calculation similar to that done above gives that

$$\widehat{\mathcal{C}}^f_{Z^\times,B^\times}M_{z^{-1}}\mathcal{O}^b_{C,A}x = Z^\times\mathcal{C}^f_{Z^\times,B^\times}\mathcal{O}^b_{C,A}x.$$

Thus (5.15) can be written more succinctly as

$$\begin{bmatrix} -Z^\times\widehat{\mathcal{C}}^f_{Z^\times,B^\times}\widehat{\mathcal{O}}^b_{C,A} & \Gamma^\times & -B^\times \end{bmatrix}\begin{bmatrix} x \\ x^\times \\ u \end{bmatrix} = 0. \tag{5.16}$$

Combining (5.14) and (5.16), we conclude that $h$ *of the form* (5.12) *is in* $z\mathcal{M}^\times\cap\mathcal{M}$ *exactly when* $\left[\begin{smallmatrix} x \\ x^\times \\ u \end{smallmatrix}\right]$ *is in* Ker $\Psi$.

By assumption the input pairs $(C, A)$ and $(C^\times, A^\times)$ are observable, and hence the correspondence

$$\begin{bmatrix} x \\ x^\times \\ u \end{bmatrix} \mapsto C(zI-A)^{-1}x + zC^\times(I-zA^\times)^{-1}x^\times + u$$

between Ker $\Psi$ and $\mathcal{L}$ is bijective. Under the assumption that $(\mathcal{M}_{\mathfrak{S}}, \mathcal{M}^\times_{\mathfrak{S}^\times})$ is a dual shift-invariant pair, the proof of Theorem 3.4 tells us that $\mathcal{L}$ and $\mathcal{U}$ are isomorphic. Hence we may choose an isomorphism between $\mathcal{L}$ and Ker $\Psi$ of the form (5.5). Then, according to the prescription in the proof of Theorem 3.4, $W(z)$ as in (5.8) gives a Beurling-Lax representation for the dual shift-invariant pair $(\mathcal{M}, \mathcal{M}^\times)$ as in (5.2).

As was shown in the proof of Theorem 3.4, $W^{*-1}$ arises as the Beurling-Lax representer for the dual shift-invariant pair $((\mathcal{M}^\times)^\perp, \mathcal{M}^\times)$. By the formula (6.2) proved below, we know that

$$(\mathcal{M}^\times)^\perp = \mathcal{M}_{(\mathfrak{S}^\times)^\perp}, \qquad \mathcal{M}^\perp = \mathcal{M}^\times_{\mathfrak{S}^\perp}$$

where the admissible Sylvester sets $(\mathfrak{S}^\times)^\perp$ and $\mathfrak{S}^\perp$ are given by

$$(\mathfrak{S}^\times)^\perp = (B^{\times *}, Z^{\times *}, A^{\times *}, C^{\times *}, -\Gamma^{\times *}), \qquad \mathfrak{S}^\perp = (B^*, Z^*, A^*, C^* - \Gamma^*). \tag{5.17}$$

Let us denote the constraining operator given by (5.3) but with $((\mathfrak{S}^\times)^\perp, \mathfrak{S}^\perp)$ in place of $(\mathfrak{S}, \mathfrak{S}^\times)$ by $\Psi_\perp$. Then careful substitution gives

$$\Psi_\perp = \begin{bmatrix} -\Gamma^{\times *} & -A^{\times *}\widehat{\mathcal{C}}^b_{A^{\times *},C^{\times *}}\widehat{\mathcal{O}}^f_{B^*,Z^*} & -C^{\times *} \\ -A^*\widehat{\mathcal{C}}^f_{A^*,C^*}\widehat{\mathcal{O}}^b_{B^{\times *},Z^{\times *}} & -\Gamma^* & -C^* \end{bmatrix}$$

as an operator from $\mathcal{D}(\Gamma^{\times *})\oplus\mathcal{D}(\Gamma^*)\oplus\mathcal{U}$ into $\mathcal{X}^\times_{\mathcal{P}}\oplus\mathcal{X}_{\mathcal{P}}$. We follow the recipe arrived at in the first part of the proof with this new data set and choose an isomorphism of the form

$$\begin{bmatrix}(C_+^\times)^* \\ (C_-^\times)^* \\ (D^\times)^*\end{bmatrix} : \mathcal{U} \to \text{Ker } \Psi_\perp \subset \begin{bmatrix}\mathcal{D}(\Gamma^{\times *}) \\ \mathcal{D}(\Gamma^*) \\ \mathcal{U}\end{bmatrix}.$$

Then we arrive at a Beurling-Lax representer $W_\perp$ for $((\mathcal{M}^\times)^\perp, \mathcal{M}^\perp)$ of the form

$$W_\perp(z) = B^{\times *}(zI - Z^{\times *})^{-1}(C_+^\times)^* + (D^\times)^* + zB^*(I - zZ^*)^{-1}(C_-^\times)^*.$$

Therefore we find that

$$W_\perp(1/\overline{z})^* = zC_+^\times(I - zZ^\times)^{-1}B^\times + D^\times + C_-^\times(zI - Z)^{-1}B \tag{5.18}$$

where the block row matrix

$$\begin{bmatrix}C_+^\times & C_-^\times & D^\times\end{bmatrix} = \begin{bmatrix}(C_+^\times)^* \\ (C_-^\times)^* \\ (D^\times)^*\end{bmatrix}^*$$

is characterized as a bounded linear operator from $\mathcal{X}_\mathcal{P}^\times \oplus \mathcal{X}_\mathcal{P} \oplus \mathcal{U}$ onto $\mathcal{U}$ with kernel equal to $(\text{Ker } \Psi_\perp)^\perp = \text{Ran } \Psi_\perp^*$. Finally we note that $\Psi_\perp^* = \widetilde{\Psi}$ where $\widetilde{\Psi}$ is given by (5.4). Thus $W(1/\overline{z})^{*-1}$, apart from the constant factor of $L^{-1}$ in the front, is given by the expression (5.9).

From the uniqueness part of Theorem 3.4, we know that there is a constant invertible operator $L$ on $\mathcal{U}$ so that

$$W_\perp(1/\overline{z})^{*-1}L = W(z)$$

where $W(z)$ is given by (5.8). The constant operator $L$ is then determined by

$$L = W_\perp(1/\overline{z})^* W(z). \tag{5.19}$$

To find $L$ (denoted for the moment $L(z)$ but it eventually will turn out to be a constant function), we compute

$$\begin{aligned} L(z) &= (zC_+^\times(I - zZ^\times)^{-1}B^\times + C_-^\times(zI - Z)^{-1}B + D^\times) \\ &\quad \cdot (C(zI - A)^{-1}B_- + zC^\times(I - zA^\times)^{-1}B_+ + D) \\ &= \begin{bmatrix}C_+^\times & C_-^\times & D^\times\end{bmatrix} Q(z) \begin{bmatrix}B_- \\ B_+ \\ D\end{bmatrix} \end{aligned}$$

where we have set $Q(z)$ equal to

$$\begin{aligned} Q(z) &= \begin{bmatrix}Q_{11}(z) & Q_{12}(z) & Q_{13}(z) \\ Q_{21}(z) & Q_{22}(z) & Q_{23}(z) \\ Q_{31}(z) & Q_{32}(z) & Q_{33}(z)\end{bmatrix} \\ &:= \begin{bmatrix}z(I - zZ^\times)^{-1}B^\times \\ z^{-1}(I - z^{-1}Z)^{-1}B \\ I\end{bmatrix} \begin{bmatrix}z^{-1}C(I - z^{-1}A)^{-1} & zC^\times(I - zA^\times)^{-1} & I\end{bmatrix}. \end{aligned} \tag{5.20}$$

Since $(Z^\times, B^\times)$ and $(Z, B)$ are exactly controllable, we see that the left factor on the right hand side of (5.20) is in $L^2_{\mathcal{L}(\mathcal{U}, \mathcal{X}_{\mathcal{Z}}^\times \oplus \mathcal{X}_{\mathcal{Z}} \oplus \mathcal{U})}$; similarly, since $(C, A)$ and $(C^\times, A^\times)$ are exactly observable by assumption, we see that the right factor on the right hand side of (5.20) is in $L^2_{\mathcal{L}(\mathcal{X}_{\mathcal{P}} \oplus \mathcal{X}_{\mathcal{P}}^\times \oplus \mathcal{U}, \mathcal{U})}$. Hence the product $Q(z)$ has each block entry in the appropriate operator-valued $L^1$-space by Proposition 2.1. We compute

$$\begin{aligned} Q_{11}(z) &= z(I - zZ^\times)^{-1} B^\times \cdot z^{-1} C (I - z^{-1} A)^{-1} \\ &= \sum_{i,j \in \mathbb{Z}_+} Z^{\times i} B^\times C A^j z^{i+1} z^{-j-1} \end{aligned}$$

and hence the Fourier coefficients $Q_{11}(z) = \sum_{n \in \mathbb{Z}} [Q_{11}]_n z^n$ for $Q_{11}$ are given by

$$\begin{aligned} [Q_{11}]_n &= \sum_{i,j \in \mathbb{Z}_+ :\, i-j=n} Z^{\times i} B^\times C A^j \\ &= \begin{cases} \sum_{n=0}^\infty Z^{\times n} B^\times C A^n = \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} & \text{if } n = 0, \\ \sum_{j=0}^n Z^{\times n+j} B^\times C A^j = Z^{\times n} \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} & \text{if } n \geq 1, \\ \sum_{i=0}^\infty Z^{\times i} B^\times C A^{i-n} = \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} A^{-n} & \text{if } n \leq -1. \end{cases} \end{aligned}$$

Thus $Q_{11}(z)$ is given by

$$\begin{aligned} Q_{11}(z) &= \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} + \sum_{n=1}^\infty Z^\times \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} z^n + \sum_{n=1}^\infty \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} A^n z^{-n} \\ &= \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} + z(I - zZ^\times)^{-1} Z^\times \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} \\ &\quad + \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} A z^{-1} (I - z^{-1} A)^{-1}. \end{aligned} \tag{5.21}$$

Similarly

$$Q_{12}(z) = z(I - zZ^\times)^{-1} B^\times \cdot z C^\times (I - zA^\times)^{-1} = \sum_{i,j \in \mathbb{Z}_+} Z^{\times i} B^\times C^\times A^{\times j} z^{i+j+2}.$$

Thus $[Q_{12}]_n = 0$ for $n \leq 1$. If we restrict all operators to $\mathcal{D}(\Gamma)$ and use the Sylvester equation (4.1) associated with $\mathfrak{S}^\times$, we see that, for $n \geq 2$,

$$\begin{aligned} [Q_{12}]_n &= \sum_{i=0}^{n-2} Z^{\times i} B^\times C^\times A^{\times n-2-i} \\ &= \sum_{i=0}^{n-2} Z^{\times i} (\Gamma^\times A^\times - Z^\times \Gamma^\times) A^{\times n-2-i} \\ &= \sum_{i=0}^{n-2} \left[ Z^{\times i} \Gamma^\times A^{\times n-1-i} - Z^{\times i+1} \Gamma^\times A^{\times n-2-i} \right] \\ &= \Gamma^\times A^{\times n-1} - Z^{\times n-1} \Gamma^\times \end{aligned}$$

and hence

$$\begin{aligned} Q_{12}(z) &= \sum_{n=2}^{\infty} \Gamma^{\times} A^{\times n-1} z^{n} - \sum_{n=2}^{\infty} Z^{\times n-1} \Gamma^{\times} z^{n} \\ &= \sum_{n=1}^{\infty} \Gamma^{\times} A^{\times n-1} z^{n} - \sum_{n=1}^{\infty} Z^{\times n-1} \Gamma^{\times} z^{n} \\ &= \Gamma^{\times} \cdot z(I - zA^{\times})^{-1} - z(I - zZ^{\times})^{-1}\Gamma^{\times}. \end{aligned} \tag{5.22}$$

Similar computations or direct copying from the definition (5.20) give us the remaining entries of $Q(z)$:

$$\begin{aligned} &Q_{13}(z) = z(I - zZ^{\times})^{-1}B^{\times}, \qquad Q_{21}(z) = \Gamma \cdot z(I - zA)^{-1} - z(I - zZ)^{-1}\Gamma, \\ &Q_{22}(z) = \widehat{\mathcal{C}}^{b}_{Z,B}\widehat{\mathcal{O}}^{f}_{C^{\times},A^{\times}} + \widehat{\mathcal{C}}^{b}_{Z,B}\widehat{\mathcal{O}}^{f}_{C^{\times},A^{\times}} A^{\times} \cdot z(I - zA^{\times})^{-1} \\ &\qquad\quad + z^{-1}(I - z^{-1}Z)^{-1} Z \widehat{\mathcal{C}}^{b}_{Z,B}\widehat{\mathcal{O}}^{f}_{C^{\times},A^{\times}}, \\ &Q_{23}(z) = z^{-1}(I - z^{-1}Z)^{-1}B, \qquad Q_{31}(z) = z^{-1}C(I - z^{-1}A)^{-1}, \\ &Q_{32}(z) = zC^{\times}(I - zA^{\times})^{-1}, \qquad Q_{33}(z) = I. \end{aligned} \tag{5.23}$$

We emphasize that we view expressions of the sort

$$z^{-1}(I - z^{-1}Z)^{-1}, \qquad z(I - zZ^{\times})^{-1}, \qquad z^{-1}(I - z^{-1}A)^{-1}, \qquad z(I - zA^{\times})^{-1}$$

as formal Fourier series in powers of $z$ and $z^{-1}$; only after performing the operations indicated in (5.24) at the level of formal Fourier series do we arrive at a formal Fourier series which is guaranteed to be the Fourier series of the weak $H^1$-function $Q(z)$.

We reorganize this information (5.21), (5.22), (5.23) in the form

$$\begin{aligned} Q(z) = &- \begin{bmatrix} z(I - zZ^{\times})^{-1} & 0 \\ 0 & (zI - Z)^{-1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \Psi \\ &- \widetilde{\Psi} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} z^{-1}(I - z^{-1}A)^{-1} & 0 & 0 \\ 0 & z(I - zA^{\times})^{-1} & 0 \end{bmatrix} \\ &+ \begin{bmatrix} \widehat{\mathcal{C}}^{f}_{Z^{\times},B^{\times}}\widehat{\mathcal{O}}^{b}_{C,A} & 0 & 0 \\ 0 & \widehat{\mathcal{C}}^{b}_{Z,B}\widehat{\mathcal{O}}^{f}_{C^{\times},A^{\times}} & 0 \\ 0 & 0 & I \end{bmatrix}. \end{aligned} \tag{5.24}$$

Since

$$\Psi \begin{bmatrix} B_{-} \\ B_{+} \\ D \end{bmatrix} = 0 \text{ and } \begin{bmatrix} C_{+}^{\times} & C_{-}^{\times} & D^{\times} \end{bmatrix} \widetilde{\Psi} = 0$$

by construction, we see that $L(z)$ collapses to the $L$ given by (5.7) since

$$\begin{aligned} L(z) &= \begin{bmatrix} C_+^\times & C_-^\times & D^\times \end{bmatrix} Q(z) \begin{bmatrix} B_- \\ B_+ \\ D \end{bmatrix} \\ &= \begin{bmatrix} C_+^\times & C_-^\times & D^\times \end{bmatrix} \begin{bmatrix} \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} & 0 & 0 \\ 0 & \widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{C^\times, A^\times} & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} B_- \\ B_+ \\ D \end{bmatrix} \\ &= C_+^\times \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} B_- + C_-^\times \widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{C^\times, A^\times} B_+ + D^\times D = L. \end{aligned}$$

From the general theory from the proof of Theorem 3.4, we know that $L$ is invertible and the formula (5.9) follows as well.

To verify that (5.8) and (5.9) provide a strongly bi-dichotomous realization for $W$, it remains only to check that $(A, B_-)$ and $(A^\times, B_+)$ are controllable and that $(C_-^\times, Z)$ and $(C_+^\times, Z^\times)$ are observable. We check only that $(A, B_-)$ is controllable as the others are similar. Since $L^2\text{-clos}\, W H^\infty_{\mathcal{U}} = \mathcal{M} = \mathcal{M}_{\mathfrak{S}}$, it follows from the proof of Theorem 4.5 that it must be the case that

$$\operatorname{Ran} \widehat{\mathcal{C}}^b_{A,B_-}|_{\mathcal{P}_{\mathcal{U}}} \text{ is dense in } \mathcal{D}(\Gamma).$$

As $\mathcal{D}(\Gamma)$ is dense in $\mathcal{X}_{\mathcal{P}}$ we conclude that $\operatorname{Ran} \widehat{\mathcal{C}}^b_{A,B_-}|_{\mathcal{P}_{\mathcal{U}}}$ is dense in $\mathcal{X}_{\mathcal{P}}$, i.e., the input pair $(A, B_-)$ is controllable. This completes the proof of Theorem 5.1. □

*Remark* 5.2. As a corollary of the proof of Theorem 3.4 and part (3) of Theorem 4.2, we see that $\Gamma_{\mathfrak{S},\mathfrak{S}^\times}$ invertible implies that dim Ker $\Psi = \dim \mathcal{U}$, i.e., after some transformations to write the result in the starkest terms, under the assumption that $\mathfrak{S} = (C, A, Z, B, \Gamma)$ and $\mathfrak{S}^\times = (C^\times, A^\times, Z^\times, B^\times, \Gamma^\times)$ are both admissible Sylvester data sets, we have

$$\begin{aligned} &\begin{bmatrix} \Gamma & \widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{C^\times, A^\times} \\ \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} & \Gamma^\times \end{bmatrix} \text{ invertible} \\ &\qquad \Longrightarrow \dim \operatorname{Ker} \begin{bmatrix} \Gamma & Z \widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{C^\times, A^\times} & -B \\ Z^\times \widehat{\mathcal{C}}^f_{Z^\times, B^\times} \widehat{\mathcal{O}}^b_{C,A} & \Gamma^\times & B^\times \end{bmatrix} = \dim \mathcal{U}. \end{aligned}$$

We do not know how to prove this directly except in some special cases. One such special case is the case where both $A^\times$ and $Z^\times$ are invertible. In this case one can use the identity

$$\Psi \cdot \begin{bmatrix} I & 0 \\ 0 & -A^\times \\ 0 & -C^\times \end{bmatrix} = \begin{bmatrix} 0 & I \\ Z^\times & 0 \end{bmatrix} \Gamma_{\mathfrak{S},\mathfrak{S}^\times}$$

to see that Ker $\Psi$ is isomorphic to $\mathcal{U}$ when $\Gamma_{\mathfrak{S},\mathfrak{S}^\times}$ is invertible. In this case one choice of $\left[\begin{smallmatrix} B_- \\ B_+ \\ D \end{smallmatrix}\right]$ is

$$\begin{bmatrix} B_- \\ B_+ \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & -A^\times \\ 0 & -C^\times \end{bmatrix} \left(\Gamma_{\mathfrak{S},\mathfrak{S}^\times}\right)^{-1} \begin{bmatrix} (Z^\times)^{-1}B^\times \\ B \end{bmatrix}. \tag{5.25}$$

In this case we may write

$$\begin{aligned} zC^\times(I - zA^\times)^{-1} &= -zC^\times A^{\times-1}(zI - A^{\times-1})^{-1} \\ &= -CA^{\times-1} - CA^{\times-1}(zI - A^{\times-1})^{-1}A^{\times-1}. \end{aligned}$$

Then we have

$$\begin{aligned} W(z) &= \begin{bmatrix} C(zI - A)^{-1} & zC^\times(I - zA^\times)^{-1} \end{bmatrix} \begin{bmatrix} B_- \\ B_+ \end{bmatrix} + D \\ &= \begin{bmatrix} C & C^\times A^{\times-1} \end{bmatrix} \begin{bmatrix} (zI - A)^{-1} & 0 \\ 0 & (zI - A^{\times-1})^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -A^{\times-1} \end{bmatrix} \begin{bmatrix} B_- \\ B_+ \end{bmatrix} \\ &\quad - C^\times A^{\times-1}B_+ + D. \end{aligned}$$

If we now use (5.25) to plug in for $B_-, B_+, D$ we get

$$\begin{aligned} W(z) &= \begin{bmatrix} C & C^\times A^{\times-1} \end{bmatrix} \begin{bmatrix} (zI - A)^{-1} & 0 \\ 0 & (zI - A^{\times-1})^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -A^{\times-1} \end{bmatrix} \cdot \\ &\qquad \cdot \begin{bmatrix} I & 0 \\ 0 & -A^\times \end{bmatrix} (\Gamma_{\mathfrak{S},\mathfrak{S}^\times})^{-1} \begin{bmatrix} Z^{\times-1}B^\times \\ B \end{bmatrix} \\ &\quad - C^\times A^{\times-1} \begin{bmatrix} 0 & -A^\times \end{bmatrix} (\Gamma_{\mathfrak{S},\mathfrak{S}^\times})^{-1} \begin{bmatrix} Z^{\times-1}B^\times \\ B \end{bmatrix} \\ &\quad + \left( I + \begin{bmatrix} 0 & -C^\times \end{bmatrix} (\Gamma_{\mathfrak{S},\mathfrak{S}^\times})^{-1} \begin{bmatrix} Z^{\times-1}B^\times \\ B \end{bmatrix} \right) \\ &= \begin{bmatrix} C & C^\times A^{\times-1} \end{bmatrix} \begin{bmatrix} (zI - A)^{-1} & 0 \\ 0 & (zI - A^{\times-1})^{-1} \end{bmatrix} (\Gamma_{\mathfrak{S},\mathfrak{S}^\times})^{-1} \begin{bmatrix} Z^{\times-1}B^\times \\ B \end{bmatrix} + I. \end{aligned}$$

This is the form of the solution with value at infinity equal to $I$ given in [19, Theorem 4.3.1] for the rational case.

Similarly, it would complete the theory to have a direct proof of the fact that invertibility of the coupling matrix $\Gamma_{\mathfrak{S},\mathfrak{S}^\times}$ implies the invertibility of the operator $L$ (5.7).

*Remark* 5.3. We can now give an alternative proof of the result of Theorem 4.4 concerning pseudocontinuation of bounded type as follows. Suppose that $W \in L^2_{\mathcal{L}(\mathcal{U})}$ is $L^2$-regular and that $\dim \mathcal{U} < \infty$. By Theorem 5.1, $W$ has a strongly bi-dichotomous realization (5.8), (5.9) arising from a pair of admissible Sylvester data sets

$$\mathfrak{S} = (C, A, Z, B, \Gamma), \qquad \mathfrak{S}^\times = (C^\times, A^\times, Z^\times, B^\times, \Gamma^\times)$$

for which

$$L^2\text{-clos}\, W H^\infty_{\mathcal{U}} = \mathcal{M}_{\mathfrak{S}}, \qquad L^2\text{-clos}\, W \overline{H^\infty_{\mathcal{U},0}} = \mathcal{M}^\times_{\mathfrak{S}^\times}.$$

By the proof of Theorem 4.2, we may assume without loss of generality that $\mathfrak{S}$ and $\mathfrak{S}^\times$ have the model form

$$\mathcal{X}_{\mathcal{P}} = L^2\text{-clos}\, P_{H^{2\perp}_{\mathcal{U}}}\mathcal{M}, \qquad \mathcal{X}_{\mathcal{Z}} = H^2_{\mathcal{U}} \ominus \left(\mathcal{M} \cap H^2_{\mathcal{U}}\right),$$
$$\mathcal{X}^\times_{\mathcal{P}} = L^2\text{-clos}\, P_{H^2_{\mathcal{U}}}\mathcal{M}, \qquad \mathcal{X}^\times_{\mathcal{Z}} = H^{2\perp}_{\mathcal{U}} \ominus \left(M^\times \cap H^{2\perp}_{\mathcal{U}}\right)$$

with

$$A = P_{\mathcal{X}_{\mathcal{P}}} M_z|_{\mathcal{X}_{\mathcal{P}}}, \qquad Z = P_{\mathcal{X}_{\mathcal{P}}} M_z|_{\mathcal{X}_{\mathcal{Z}}}$$
$$A^\times = P_{\mathcal{X}^\times_{\mathcal{P}}} M_{z^{-1}}|_{\mathcal{X}^\times_{\mathcal{P}}}, \qquad Z^\times = P_{\mathcal{X}^\times_{\mathcal{Z}}} M_{z^{-1}}|_{\mathcal{X}^\times_{\mathcal{Z}}}.$$

As we have seen via Lemma 4.3, the fact that $\mathcal{M}$ is full-range simply-invariant for $M_z$ and that $\mathcal{M}^\times$ is full-range simply-invariant for $M_{z^{-1}}$ implies hat $A$,$Z$,$A^\times$, and $Z^\times$ are all strongly bi-stable, i.e., are in the Sz.-Nagy-Foias class $C_{00}$ with characteristic function equal to a matrix-valued (since $\dim \mathcal{U} < \infty$) two-sided inner function. As all these operators are contractions, certainly all have spectra in the closed unit disk $\overline{\mathbb{D}}$. In case $\dim \mathcal{U} < \infty$, one can say more: *the intersection of the spectrum with the open unit disk consists only of eigenvalues of finite multiplicity having no accumulation point in* $\mathbb{D}$ *and consequently each of the resolvents* $(zI - A)^{-1}$, $(zI-Z)^{-1}$, $(zI-A^\times)^{-1}$ *and* $(zI-Z^\times)^{-1}$ *is meromorphic on* $\mathbb{D}$ (see e.g. [38, Section VI.4]). We then see that $z(I-zA^\times)^{-1} = (z^{-1}I-A^\times)^{-1}$ and $z(I-zZ^\times)^{-1} = (z^{-1}I - Z^\times)^{-1}$ are meromorphic on $\mathbb{D}$. As $A$, $Z$, $A^\times$ and $Z^\times$ are all contractions, by the Neumann series expansion we read off that $(zI - A)^{-1}$ and $(zI - Z)^{-1}$ are analytic on $\mathbb{D}_e$ and that $z(I - zA^\times)^{-1}$ and $z(I - zZ^\times)^{-1}$ are analytic on $\mathbb{D}_e$. Then from the formulas for $W$ and $W^{-1}$ (5.8) and (5.9) we read off that $W$ and $W^{-1}$ have meromorphic pseudocontinuation to both $\mathbb{D}$ and $\mathbb{D}_e$. Moreover, since $\dim \mathcal{U} < \infty$, it is possible to argue that these pseudocontinuations are of bounded type, but we omit the details concerning this point. By the observability/controllability properties of the realizations (5.8) and (5.9), one can say more: the poles of $W$ in $\mathbb{D}$ consist exactly of the eigenvalues of $A$ in $\mathbb{D}$, the poles of $W$ in $\mathbb{D}_e$ consist exactly of the reflections $1/\lambda$ of the eigenvalues $\lambda$ of $A^\times$ in $\mathbb{D}$ (where we interpret $\infty = 1/0$), the poles of $W^{-1}$ in $\mathbb{D}$ consist exactly of the eigenvalues of $Z$ inside $\mathbb{D}$, and the poles of $W^{-1}$ in $\mathbb{D}_e$ consist exactly of the reflections $1/\lambda$ of the eigenvalues $\lambda$ of $Z^\times$ in $\mathbb{D}$. In this way we arrive at a realization-theoretic proof of the pseudocontinuation part of Theorem 4.4.

*Remark* 5.4. It is of interest that strongly bi-dichotomous realizations for a given $L^2$-regular function $W$ are unique up to a state-space similarity, i.e., *given a strongly bi-dichotomous realization* (2.18) *and* (2.19) *together with a second strongly bi-dichotomous realization for* $W$

$$W(z) = zC'_+(I_{\mathcal{X}'_+} - zA'_+)^{-1}B'_+ + D' + C'_-(zI_{\mathcal{X}'_-} - A'_-)^{-1}B'_- \tag{5.26}$$

$$W(z)^{-1} = zC'^\times_+(I_{\mathcal{X}'^\times_+} - zA'^\times_+)^{-1}B'^\times_+ + D'^\times + C'^\times_-(zI_{\mathcal{X}'^\times_-} - A'^\times_-)^{-1}B'^\times_-, \tag{5.27}$$

*then there are bounded and invertible similarity transforms*

$$S_+ \in \mathcal{L}(\mathcal{X}_+, \mathcal{X}'_+), \qquad S_- \in \mathcal{L}(\mathcal{X}_-, \mathcal{X}'_-),$$
$$S_+^\times \in \mathcal{L}(\mathcal{X}_+^\times, \mathcal{X}_+'^\times) \qquad S_-^\times \in \mathcal{L}(\mathcal{X}_-^\times, \mathcal{X}_-'^\times)$$

*so that*

$$\begin{aligned}
&C'_+ = C_+ S_+^{-1}, && A'_+ = S_+ A_+ S_+^{-1}, && B'_+ = S_+ B_+,\\
&C'_- = C_- S_-^{-1}, && A'_- = S_- A_- S_-^{-1}, && B'_- = S_- B_-,\\
&C_+'^\times = C_+^\times S_+^{\times -1}, && A_+^{\times\prime} = S_+^\times A_+ S_+^{\times -1}, && B_+'^\times = S_+^\times B_+,\\
&C_-'^\times = C_-^\times S_-^{\times -1}, && A_-'^\times = S_-^\times A_-^\times S_-^{\times -1}, && B_-'^\times = S_-^\times B_-.
\end{aligned} \tag{5.28}$$

One way to prove this is to show that any strongly bi-dichotomous realization is similar to the model bi-dichotomous realization coming from Theorem 5.1 with Sylvester data sets $(\mathfrak{S}, \mathfrak{S}^\times)$ taken to be the model Sylvester data sets arising from the dual shift-invariant pair $(L^2\text{-clos}\, W H^\infty_{\mathcal{U}}, L^2\text{-clos}\, W \overline{H^\infty_{\mathcal{U},0}})$. However one can proceed more directly as follows. By equating Fourier coefficients in the two supposed realizations for $W(z)$, we get the moment equalities

$$C'_+ A_+'^n B'_+ = C_+ A_+^n B_+ \text{ for } n \in \mathbb{Z}_+,\ D' = D,\ C'_- A_-'^n B'_- = C_- A_-^n B_- \text{ for } n \in \mathbb{Z}_+.$$

To get the first set of equalities in (5.28), it suffices to find a bounded, invertible similarity $T = S_+^{*-1} \colon \mathcal{X}_+ \to \mathcal{X}'_+$ so that

$$C_+'^* = T C_+^*, \qquad A_+'^* = T A_+ T^{-1}, \qquad B'^* = B^* T^{-1}. \tag{5.29}$$

It is then natural to define $T$ so that

$$T \colon A_+^{*n} C_+^* u \mapsto A_+'^* C_+'^*.$$

which then forces

$$T \mathcal{C}_{A_+^*, C_+^*} \{u\}_{n \in \mathbb{Z}_+} = \mathcal{C}_{A_+'^*, C_+'^*} \{u\}_{n \in \mathbb{Z}_+}.$$

Since as part of a strongly bi-dichotomous realization we know that $(C_+, A_+)$ is exactly observable, it follows that

$$T = \mathcal{C}_{A_+'^*, C_+'^*} (\mathcal{C}_{A_+^*, C_+^*})^{-1} = (\mathcal{O}_{C'_+, A'_+})^* (\mathcal{O}_{C_+, A_+})^{*-1}$$

is bounded and invertible. It is now straightforward to check that $T$ meets all the intertwining conditions (5.29). The other intertwining relations in (5.28) can be solved in a similar way using the exact observability of $(C_-, A_-)$ and $(C'_-, A'_-)$ and the exact controllability of $(A_+^\times, B_+^\times)$, $(A_+'^\times, B_+'^\times)$, $(A_-^\times, B_-^\times)$, and $(A_-'^\times, B_-'^\times)$.

In a similar vein, one can show that *two admissible Sylvester data sets*

$$\mathfrak{S} = (C, A, Z, B, \Gamma), \qquad \mathfrak{S}' = (C', A', Z', B', \Gamma')$$

*give rise to the same full-range $M_z$-simply-invariant subspace $\mathcal{M}_{\mathfrak{S}} = \mathcal{M}_{\mathfrak{S}'}$ if and only if $\mathfrak{S}$ and $\mathfrak{S}'$ are* similar *in the sense that there are bounded, invertible similarity transformations*

$$S \colon \mathcal{X}_{\mathcal{P}} \to \mathcal{X}'_{\mathcal{P}}, \qquad T \colon \mathcal{X}_{\mathcal{Z}} \to \mathcal{X}'_{\mathcal{Z}}$$

*so that*

$$C' = CS^{-1}, \qquad A' = SAS^{-1}, \qquad Z' = TZT^{-1}, \qquad B' = TB, \qquad \Gamma' = T\Gamma S^{-1}.$$

A similar statement holds for two admissible Sylvester data sets giving rise to the same full-range $M_{z^{-1}}$-simply-invariant subspace $\mathcal{M}^{\times}_{\mathfrak{S}} = \mathcal{M}^{\times}_{\mathfrak{S}'}$.

## 6. $J$-orthogonal dual pairs of shift-invariant subspaces and strongly regular $J$-unitary functions

We now assume that $J$ is an operator on $\mathcal{U}$ such that $J = J^* = J^{-1}$. We identify $J$ with $J \otimes I_{L^2}$ (the operator of multiplication by the constant function $J$ on $L^2_{\mathcal{U}}$). We say that the $\mathcal{L}(\mathcal{U})$-valued function $\Theta(z)$ is a *$J$-unitary function* if $\Theta(z)^* J \Theta(z) = J$ for almost all $z \in \mathbb{T}$. We are interested in adapting the results of the previous sections to the case where $W = \Theta$ is a $J$-unitary function in $L^2_{\mathcal{L}(\mathcal{U})}$.

From the $J$-unitary property of $\Theta(z)$ for $z \in \mathbb{T}$, it follows that $\Theta(z)^{-1} = J\Theta(z)^* J$ for almost all $z \in \mathbb{T}$, and therefore we also have $\Theta^{-1} \in L^2_{\mathcal{L}(\mathcal{U})}$. Following the terminology of Definition 3.3, we say that the $J$-unitary function $\Theta$ is *$L^2$-regular $J$-unitary* if the operator

$$M_\Theta P_{H^2_{\mathcal{U}}} M_\Theta^{-1} \colon L^\infty_{\mathcal{U}} \to L^1_{\mathcal{U}}$$

extends to define a bounded operator from $L^2_{\mathcal{U}}$ to itself. From the identity $\Theta^{-1} = J\Theta^* J$ we see that the $J$-unitary function $\Theta$ is $L^2$-regular exactly when the operator $M_\Theta J P_{H^2_{\mathcal{U}}} M^*_\Theta J$ extends to define a bounded operator on $L^2_{\mathcal{U}}$. This operator in turn is a $J$-self-adjoint projection operator onto the subspace $\mathcal{M} := L^2\text{-clos}\, \Theta H^\infty_{\mathcal{U}}$. Thus the $J$-unitary function $\Theta$ being $L^2$-regular means simply that the subspace $\mathcal{M} := L^2\text{-clos}\, \Theta H^2_{\mathcal{U}}$ is the range of a bounded $J$-self-adjoint projection operator, i.e., that $\mathcal{M}$ is *regular* as a subspace of $L^2_{\mathcal{U}}$ considered as a Kreĭn space with indefinite inner product $\langle \cdot, \cdot \rangle_J$ induced by $J$:

$$\langle f, g \rangle_J := \langle Jf, g \rangle_{L^2_{\mathcal{U}}} = \frac{1}{2\pi} \int_{\mathbb{T}} \langle Jf(\zeta), g(\zeta) \rangle_{\mathcal{U}} \; |d\zeta|.$$

If $\Theta$ is $J$-unitary and $\mathcal{M} = L^2\text{-clos}\, \Theta H^\infty_{\mathcal{U}}$, then one can check that

$$\begin{aligned} \mathcal{M}^{\times} &:= L^2\text{-clos}\, \Theta \overline{H^\infty_{\mathcal{U}}} \\ &= L^2\text{-clos}\, J\Theta^{*-1} J \overline{H^\infty_{\mathcal{U}}} \\ &= L^2\text{-clos}\, J\Theta^{*-1} \overline{H^\infty_{\mathcal{U}}} \\ &= \left( L^2\text{-clos}\, \Theta H^\infty_{\mathcal{U}} \right)^{\perp J} \\ &= \mathcal{M}^{\perp J}. \end{aligned}$$

With these observations, Theorem 3.4 assumes the following form.

**Theorem 6.1.** *(See* [20, 21]*.) Suppose that $\mathcal{M}$ is a subspace of $L^2_{\mathcal{U}}$. Then the following are equivalent.*

1. *$\mathcal{M}$ is a full-range, simply-invariant subspace for $M_z$ which is regular as a subspace of the Kreĭn space $(L^2_{\mathcal{U}}, \langle\cdot,\cdot\rangle_J)$.*
2. *There exists an $L^2$-regular $J$-unitary function $\Theta$ so that $\mathcal{M} = L^2\text{-clos}\,\Theta H^\infty_{\mathcal{U}}$.*

*Proof.* Suppose first that $\mathcal{M}$ is a full-range simply-invariant $J$-regular subspace This means that the pair $(\mathcal{M}, \mathcal{M}^{\perp J})$ is a dual shift-invariant pair. By Theorem 3.4 there is an $L^2$-regular $\mathcal{L}(\mathcal{U})$-valued function $W$ so that $\mathcal{M} = L^2\text{-clos}\,WH^\infty_{\mathcal{U}}$ and $\mathcal{M}^{\perp J} = L^2\text{-clos}\,W\overline{H^\infty_{\mathcal{U},0}}$. It is not difficult to check that $W' = JW^{*-1}$ also serves as a representer for $\mathcal{M}$ and $\mathcal{M}^{\perp J}$. By the uniqueness assertion in Theorem 3.4, it follows that there is an invertible constant operator $X \in \mathcal{L}(\mathcal{U})$ so that $JW^{*-1} = WX^{-1}$, i.e., so that

$$W(z)^*JW(z) = X \text{ for almost all } z \in \mathbb{T}.$$

Thus $X$ must have the same inertia as $J$, i.e., there is a constant operator $X' \in \mathcal{L}(\mathcal{U})$ so that $X = X'^*JX'$. If we then set $\Theta(z) = W(z)X'^{-1}$, then $\Theta(z)$ meets all the requirements of part (2) of Theorem 6.1. The converse direction is easily checked directly. This completes the proof of Theorem 6.1. □

**Proposition 6.2.** *Suppose that $\mathfrak{S} = (C, A, Z, B, \Gamma)$ is an admissible Sylvester data set as in Definition* 4.1 *and let $\mathcal{M}_{\mathfrak{S}}$ be the associated full-range simply-invariant subspace for $M_z$. Then*

$$\mathcal{M}_{\mathfrak{S}}^{\perp J} = \mathcal{M}^\times_{\mathfrak{S}^{\perp J}}, \qquad (\mathcal{M}^\times_{\mathfrak{S}})^{\perp J} = \mathcal{M}_{\mathfrak{S}^{\perp J}} \tag{6.1}$$

*where we have set*

$$\mathfrak{S}^{\perp J} = (JB^*, Z^*, A^*, C^*J, -\Gamma^*). \tag{6.2}$$

*Proof.* We prove only the first of the identities (6.1) as the proof of the second is completely similar.

First note that it is an easy consequence of the Definition 4.1 that $\mathfrak{S}^{\perp J}$ is again an admissible Sylvester data set whenever $\mathfrak{S}$ is an admissible Sylvester data set. Indeed, exact observability of $(C, A)$ is equivalent to exact controllability of $(A^*, C^*)$ and hence also for $(A^*, C^*J)$, exact controllability of $(Z, B)$ is equivalent to exact observability for $(B^*, Z^*)$ and hence also for $(JB^*, Z^*)$, and it is an easy (if somewhat delicate) exercise to show that

$$\begin{aligned}\Gamma Ax - Z\Gamma x = BCx \text{ for all } x \in \mathcal{D}(\Gamma) \iff \\ A^*\Gamma^* y - \Gamma^* Z^* y = C^*B^* y \text{ for all } y \in \mathcal{D}(\Gamma^*).\end{aligned}$$

Thus $\mathcal{M}^\times_{\mathfrak{S}^{\perp J}}$ is a closed $M_{z^{-1}}$-invariant subspace of $L^2_{\mathcal{U}}$ along with $\mathcal{M}$ being a closed $M_z$-invariant subspace.

Suppose now that $g + h$ is $J$-orthogonal to all of $\mathcal{M}_{\mathfrak{S}}$, where $g \in H^{2\perp}_{\mathcal{U}}$ and $h \in H^2_{\mathcal{U}}$. In particular, necessarily $h$ is $J$-orthogonal to $\mathcal{M}_{\mathfrak{S}} \cap H^2_{\mathcal{U}} = \operatorname{Ker} \widehat{\mathcal{C}}^b_{Z,B}$. As $(Z, B)$ is exactly controllable by assumption, $(\widehat{\mathcal{C}}^b_{Z,B})^* = \widehat{\mathcal{O}}^f_{B^*,Z^*}$ has closed range and $(\operatorname{Ker} \widehat{\mathcal{C}}^b_{Z,B})^\perp = \operatorname{Ran} \widehat{\mathcal{O}}^f_{B^*,Z^*}$. Hence, the fact that $h$ is $J$-orthogonal to $\mathcal{M}_{\mathfrak{S}} \cap H^2_{\mathcal{U}}$ gives us that there exists a $y \in \mathcal{X}_{\mathcal{Z}}$ so that $h = \widehat{\mathcal{O}}^f_{JB^*,Z^*} y$. Next let

$\widehat{\mathcal{O}}^b_{C,A}x + f$ be a generic element of $\mathcal{M}_{\mathfrak{S}}$. Then the $J$-orthogonality of $g+h$ to $\mathcal{M}_{\mathfrak{S}}$ gives us

$$\begin{aligned} 0 &= \langle J(g+h), \widehat{\mathcal{O}}^b_{C,A}x + f\rangle_{L^2_{\mathcal{U}}} \\ &= \langle Jg, \widehat{\mathcal{O}}^b_{C,A}x\rangle_{H^{2\perp}_{\mathcal{U}}} + \langle J\widehat{\mathcal{O}}^f_{JB^*,Z^*}y, f\rangle_{H^2_{\mathcal{U}}} \\ &= \langle \widehat{\mathcal{C}}^f_{A^*,C^*J}g, x\rangle_{\mathcal{X}_{\mathcal{P}}} + \langle y, \widehat{\mathcal{C}}^b_{Z,B}f\rangle_{\mathcal{X}_{\mathcal{Z}}} \\ &= \langle \widehat{\mathcal{C}}^f_{A^*,C^*J}g, x\rangle_{\mathcal{X}_{\mathcal{P}}} + \langle y, \Gamma x\rangle_{\mathcal{X}_{\mathcal{Z}}} \end{aligned}$$

where we used the defining condition $\widehat{\mathcal{C}}^b_{Z,B}f = \Gamma x$ for the admission of $\widehat{\mathcal{O}}^b_{C,A}x + f$ to $\mathcal{M}_{\mathfrak{S}}$ in the last step. This last identity finally shows that $y \in \mathcal{D}(\Gamma^*)$ and $\Gamma^* y = -\widehat{\mathcal{C}}^f_{A^*,C^*J}g$, i.e., that $g + \widehat{\mathcal{O}}^f_{JB^*,Z^*}y \in \mathcal{M}^\times_{\mathfrak{S}^{\perp J}}$. We conclude that $(\mathcal{M}_{\mathfrak{S}})^{\perp J} \subset \mathcal{M}^\times_{\mathfrak{S}^{\perp J}}$. The reverse containment can be checked directly. This concludes the proof of Proposition 6.2. □

*Remark* 6.3. Suppose that we are given two admissible Sylvester data sets $\mathfrak{S}$ and $\mathfrak{S}^\times$ as in (4.5). As explained in part (3) of Theorem 4.2, then the pair $(\mathcal{M}_{\mathfrak{S}}, \mathcal{M}^\times_{\mathfrak{S}^\times})$ satisfies the matching condition if and only if the coupling operator $\Gamma_{\mathfrak{S},\mathfrak{S}^\times}$ (see (4.6)) is invertible. Using the definition (6.2), substitution into (4.6) gives

$$\begin{aligned} \Gamma_{\mathfrak{S}^{\times\perp J},\mathfrak{S}^{\perp J}} &= \begin{bmatrix} -\widehat{\mathcal{C}}^f_{A^*,C^*J}\widehat{\mathcal{O}}^b_{JB^{\times *},Z^{\times *}} & \Gamma^* \\ -\Gamma^{\times *} & \widehat{\mathcal{C}}^b_{A^{\times *},C^{\times *}J}\widehat{\mathcal{O}}^f_{JB^*,Z^*} \end{bmatrix} \\ &= \begin{bmatrix} -\widehat{\mathcal{C}}^f_{A^*,C^*}\widehat{\mathcal{O}}^b_{B^{\times *},Z^{\times *}} & \Gamma^* \\ -\Gamma^{\times *} & \widehat{\mathcal{C}}^b_{A^{\times *},C^{\times *}}\widehat{\mathcal{O}}^f_{B^*,Z^*} \end{bmatrix} = \left(\Gamma_{\mathfrak{S},\mathfrak{S}^\times}\right)^*. \end{aligned}$$

In particular, $\Gamma_{\mathfrak{S}^{\times\perp J},\mathfrak{S}^{\perp J}}$ is invertible if and only if $\Gamma_{\mathfrak{S},\mathfrak{S}^\times} = \left(\Gamma_{\mathfrak{S}^{\times\perp J},\mathfrak{S}^{\perp J}}\right)^*$ is invertible. As a consequence of part (3) of Theorem 4.2, we conclude that the pair $(\mathcal{M}, \mathcal{M}^\times)$ satisfies the matching condition (4.8) if and only if the pair $((\mathcal{M}^\times)^{\perp J}, \mathcal{M}^{\perp J})$ satisfies the matching condition. This gives a concrete proof (via null-pole-data calculus) of this result which can be also proved by using general duality ideas. Note that the particular case of this result for the case $J = I_{\mathcal{U}}$ was used in the proof of Theorem 3.4.

When part (3) of Theorem 4.2 is applied to the case where $\mathcal{M}^\times = \mathcal{M}^{\perp J}$ and $\mathcal{M}$ is assumed to be in the form $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}$ for an admissible Sylvester data set, we arrive at the following corollary.

**Corollary 6.4.** *Let $\mathfrak{S} = (C, A, Z, B, \Gamma)$ be an admissible Sylvester data set. Then the associated subspace $\mathcal{M}_{\mathfrak{S}}$ is a regular subspace of the Kreĭn space $(L^2_{\mathcal{U}}, \langle\cdot,\cdot\rangle_J)$ if and only if the associated coupling operator*

$$\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}} = \begin{bmatrix} -(\widehat{\mathcal{O}}_{C,A})^* J \widehat{\mathcal{O}}_{C,A} & \Gamma^* \\ \Gamma & \widehat{\mathcal{C}}_{Z,B} J (\widehat{\mathcal{C}}_{Z,B})^* \end{bmatrix} : \begin{bmatrix} \mathcal{D}(\Gamma) \\ \mathcal{D}(\Gamma^*) \end{bmatrix} \to \begin{bmatrix} \mathcal{X}_P \\ \mathcal{X}_Z \end{bmatrix} \tag{6.3}$$

*is invertible.*

*Proof.* Simply use the result of Proposition 6.2 that $(\mathcal{M}_{\mathfrak{S}})^{\perp J} = \mathcal{M}^{\times}_{\mathfrak{S}^{\perp J}}$ and apply part (3) of Theorem 4.2. □

We are now ready to obtain a result concerning the realization of a $J$-unitary Beurling-Lax representer in terms of null-pole data.

**Theorem 6.5.** *Suppose that* $\mathfrak{S} = (C, A, Z, B, \Gamma)$ *is an admissible Sylvester data set. Then there is a* $L^2$*-regular* $J$*-unitary function* $\Theta$ *such that*

$$\mathcal{M}_{\mathfrak{S}} = L^2\text{-}clos\, \Theta \cdot H^\infty_{\mathcal{U}}$$

*if and only if the coupling operator* $\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}$ *given by* (6.3) *is invertible. If this is the case, introduce*

$$\Psi = \begin{bmatrix} \Gamma & -Z\widehat{\mathcal{C}}^b_{Z,B} J (\widehat{\mathcal{C}}^b_{Z,B})^* & -B \\ -A^* (\widehat{\mathcal{O}}^b_{C,A})^* J \widehat{\mathcal{O}}^b_{C,A} & -\Gamma^* & -C^* J \end{bmatrix},$$

*let*

$$\begin{bmatrix} B_- \\ B_+ \\ D \end{bmatrix} : \mathcal{U} \to \operatorname{Ker} \Psi \subset \begin{bmatrix} \mathcal{D}(\Gamma) \\ \mathcal{D}(\Gamma^\times) \\ \mathcal{U} \end{bmatrix}$$

*be an isomorphism between* $\mathcal{U}$ *and* Ker $\Psi$*, with the additional property that*

$$J = B_-^* (\widehat{\mathcal{O}}^b_{C,A})^* J \widehat{\mathcal{O}}^b_{C,A} B_- + B_+^* \widehat{\mathcal{C}}^b_{Z,B} J (\widehat{\mathcal{C}}^b_{Z,B})^* B_+ + D^* J D.$$

*Then a choice of* $\Theta$ *is given by*

$$\Theta(z) = C(zI - A)^{-1} B_- + D + zJB^*(I - zZ^*)^{-1} B_+. \tag{6.4}$$

*Proof.* This follows by following the recipe in Theorem 5.1 for the particular case where $\mathfrak{S}^\times = \mathfrak{S}^{\perp J}$. In this case it works out that

$$\widetilde{\Psi} = \begin{bmatrix} \Gamma^* & -\widehat{\mathcal{C}}^f_{A^*, C^* J} \widehat{\mathcal{O}}^b_{C,A} A \\ -\widehat{\mathcal{C}}^b_{Z,B} \widehat{\mathcal{O}}^f_{JB^*, Z^*} Z^* & -\Gamma \\ -JB^* & -C \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & J \end{bmatrix} \Psi^*. \tag{6.5}$$

If we use the notation $\left[\begin{smallmatrix} B'_- \\ B'_+ \\ D' \end{smallmatrix}\right]$ in place of $\left[\begin{smallmatrix} B_- \\ B_+ \\ D \end{smallmatrix}\right]$ for the column matrix constructed as in (5.5), then

$$L = B'^*_- (\widehat{\mathcal{O}}^b_{C,A})^* J \widehat{\mathcal{O}}^b_{C,A} B'_- + B'^*_+ \widehat{\mathcal{C}}^b_{Z,B} J (\widehat{\mathcal{C}}^b_{Z,B})^* B'_+ + D'^* J D'. \tag{6.6}$$

From the connection (6.5) between $\widetilde{\Psi}$ and $\Psi$, we see that, in the recipe given by Theorem 5.1, we can take

$$\begin{bmatrix} C_+^\times & C_-^\times & D^\times \end{bmatrix} = \begin{bmatrix} B_-^* & B_+^* & D^* J \end{bmatrix}.$$

Then the formula (5.8) for $W(z)$ specializes to

$$W(z) = C(zI - A)^{-1} B'_- + D' + zJB^*(I - zZ^*)^{-1} B'_+, \tag{6.7}$$

while the formula (5.18) $W_\perp(1/\overline{z})^*$ specializes to

$$W_\perp(1/\overline{z})^* = zB'^*_-(I - zA^*)^{-1} C^* J + D'^* J + B'^*_+ (zI - Z)^{-1} B. \tag{6.8}$$

Hence $W$ is a Beurling-Lax representer for the pair of subspaces $(\mathcal{M}_{\mathfrak{S}}, \mathcal{M}_{\mathfrak{S}^{\perp J}})$. If we use (6.7) to compute $W(1/\overline{z})^*$ we arrive at

$$W(1/\overline{z})^* = zB'^*_-(I - zA^*)^{-1}C^* + D'^* + B'^*_+(zI - Z)^{-1}BJ \tag{6.9}$$

and we see that

$$W_\perp(1/\overline{z})^* = W(1/\overline{z})^*J.$$

From (5.19) we see therefore that

$$W(1/\overline{z})^*JW(z) = L \tag{6.10}$$

with $L$ given by (6.6). From (6.10) we see that $L$ has a factorization $L = X^*JX$ with $X \in \mathcal{L}(\mathcal{U})$ invertible. From (6.10), it is clear that $\Theta(z) := W(z)X^{-1}$ has $J$-unitary values on $\mathbb{T}$ as wanted and that $\Theta(z)$ is given by (6.4) with

$$\begin{bmatrix} B_- \\ B_+ \\ D \end{bmatrix} = \begin{bmatrix} B'_- \\ B'_+ \\ D' \end{bmatrix} X^{-1}.$$

□

## 7. Realization theory for $J$-inner functions

In this section we assume throughout that

$$\dim \mathcal{U} < \infty. \tag{7.1}$$

Then, as a consequence of Theorem 4.4 and Remark 5.3, any $L^2$-regular function $W \in L^2_{\mathcal{L}(\mathcal{U})}$ has pseudocontinuation of bounded type to $\mathbb{D}$ and to $\mathbb{D}_e$. In particular this is true for a $L^2$-regular $J$-unitary function $\Theta$, so the values $\Theta(z)$ for $z \in \mathbb{D}$ make sense at any point of analyticity of $\Theta$ which is at all points of $\mathbb{D}$ except for the discrete set of poles in $\mathbb{D}$. We can then introduce the kernel function $K_\Theta(z,w)$ defined by

$$K_\Theta(z,w) = \frac{J - \Theta(z)J\Theta(z)^*}{1 - z\overline{w}} \tag{7.2}$$

for all pairs of points $z, w$ in $\mathbb{D}$ where $\Theta$ is analytic. We say that the $J$-unitary function $\Theta(z)$ is *$J$-inner* if the kernel $K_\Theta$ is a positive kernel on $\mathbb{D} \times \mathbb{D}$ in the sense that

$$\sum_{i,j=1}^{N} \langle K_\Theta(z_i, z_j)u_j,\, u_i\rangle \geq 0$$

for all choices of points of analyticity $z_1, \dots, z_N$ of $\Theta$ inside $\mathbb{D}$ and vectors $u_1, \dots, u_N \in \mathcal{U}$ for $N = 1, 2, \dots$. Among all $L^2$-regular $J$-unitary functions, the $L^2$-regular $J$-inner functions have special structure, as summarized in the following theorem. This result is essentially contained in [22]; for completeness we give a direct proof for our context here.

**Theorem 7.1.** *Suppose that $\Theta$ is a $L^2$-regular $J$-unitary function with associated dual pair of shift invariant subspaces*

$$\mathcal{M}_{\mathfrak{S}} = L^2\text{-}clos\,\Theta H^\infty_{\mathcal{U}}, \qquad \mathcal{M}_{\mathfrak{S}^{\perp J}} = L^2\text{-}clos\,\Theta\overline{H^\infty_{\mathcal{U},0}} = \mathcal{M}^{\perp J}$$

*arising from the admissible Sylvester data set*

$$\mathfrak{S} = (C, A, Z, B, \Gamma)$$

*as in* (4.3), (4.4) *and* (6.2). *Then the following are equivalent.*

1. *$\Theta$ is $J$-inner, i.e., the kernel* (7.2) *is a positive kernel on $\mathbb{D}\times\mathbb{D}$.*
2. *The operator $Q\colon L^2_{\mathcal{U}} \to L^2_{\mathcal{U}}$ defined by*
$$Q = JP_{H^2_{\mathcal{U}}} - JM_\Theta P_{H^2_{\mathcal{U}}} M_{\Theta^{-1}} \tag{7.3}$$
*is positive:*
$$\langle Qf,\, f\rangle_{L^2_{\mathcal{U}}} \geq 0 \text{ for all } f \in L^2_{\mathcal{U}}.$$
3. *The coupling operator $\Gamma$ is bounded and the coupling matrix $\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}$ given by* (6.3) *is positive-definite:*
$$\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}} = \begin{bmatrix} -(\widehat{\mathcal{O}}_{C,A})^* J \widehat{\mathcal{O}}_{C,A} & \Gamma^* \\ \Gamma & \widehat{\mathcal{C}}_{Z,B} J (\widehat{\mathcal{C}}_{Z,B})^* \end{bmatrix} \geq 0.$$

*Proof.* Choose a two-sided $\mathcal{L}(\mathcal{U})$-valued inner function $\psi$ so that $\psi\cdot\Theta \in H^2_{\mathcal{U}}$. Then, for $z, w$ points of analyticity for $\psi^{-1}$ in $\mathbb{D}$ and for all pairs $u, u'$ of vectors in $\mathcal{U}$, it is not difficult to see (see [22, Theorem 3.3]) that

$$\left\langle QM_{\psi^{-1}} k_{\mathrm{Sz}}(\cdot, w)\psi(w)^{*-1}u,\, M_{\psi^{-1}} k_{\mathrm{Sz}}(\cdot, z)\psi(z)^{*-1}u'\right\rangle_{L^2_{\mathcal{U}}} = \langle K_\Theta(z,w)u,\, u'\rangle_{\mathcal{U}} \tag{7.4}$$

where we use the notation $k_{\mathrm{Sz}}(z,w)$ for the Szegö kernel

$$k_{\mathrm{Sz}}(z,w) = \frac{1}{1 - z\overline{w}}.$$

From (7.4) we see immediately that (2) $\Longrightarrow$ (1). As $Q|_{\psi^{-1}\cdot H^{2\perp}_{\mathcal{U}}} = 0$, we see that (1) $\Longrightarrow$ (2) as well.

We prove (2) $\Longleftrightarrow$ (3) in detail under the additional assumption that

$$\mathcal{G}^J_{Z,B} := \widehat{\mathcal{C}}^b_{Z,B} J (\widehat{\mathcal{C}}^b_{Z,B})^* \text{ is invertible.} \tag{7.5}$$

The general case can be reduced to this case by approximating $(Z, B)$ by output pairs $(Z_n, B_n)$ which are still exactly controllable and which have the additional property that $\mathcal{G}^J_{Z_n,B_n}$ are all invertible.

We therefore assume that $\mathcal{G}^J_{Z,B}$ is invertible. In general, we have the following *general principle* for computation of $J$-orthogonal projections: *if $X\colon \mathcal{X}' \to \mathcal{N} \subset L^2_{\mathcal{U}}$ is an injective parametrization of a subspace $\mathcal{N} \subset L^2_{\mathcal{U}}$, then $\mathcal{N}$ is a regular subspace of $(L^2_{\mathcal{U}}, J)$ if and only if $X^*JX$ is invertible on $\mathcal{X}'$ and then the $J$-orthogonal projection of $L^2_{\mathcal{U}}$ onto $\mathcal{N}$ is given by*

$$P_{\mathcal{N}} = X(X^*JX)^{-1}X^*J.$$

We apply this principle to the case where $\mathcal{N} = H^2 \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$ with $X = J(\widehat{\mathcal{C}}^b_{Z,B})^*$ and $X^* J X = \mathcal{G}^J_{Z,B}$ to conclude that $H^2_{\mathcal{U}} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$ is a regular subspace of $(L^2_{\mathcal{U}}, J)$ with $J$-orthogonal projection given by

$$P_{H^2_{\mathcal{U}} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})} = J(\widehat{\mathcal{C}}^b_{Z,B})^* \left(\mathcal{G}^J_{Z,B}\right)^{-1} \widehat{\mathcal{C}}^b_{Z,B}.$$

The content of the extra assumption (7.5) is that not only $\mathcal{M}$ but also $\mathcal{M} \cap H^2_{\mathcal{U}}$, and hence furthermore $H^2_{\mathcal{U}} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$, is a regular subspace of the Kreĭn space $(L^2_{\mathcal{U}}, J)$. Then necessarily $\mathcal{M} \ominus (H^2_{\mathcal{U}} \cap \mathcal{M})$ is also a regular subspace of $(L^2_{\mathcal{U}}, J)$. To compute its $J$-orthogonal projection, we have to first find an appropriate parametrization $X\colon \mathcal{X}' \to \mathcal{M} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$. For $x \in \mathcal{D}(\Gamma)$, define $f_x \in H^2 \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$ by

$$f_x = J(\widehat{\mathcal{C}}^b_{Z,B})^* \left(\mathcal{G}^J_{Z,B}\right)^{-1} \Gamma x.$$

The point here is that $f_x$ is the unique element of $H^2_{\mathcal{U}} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$ so that $\widehat{\mathcal{O}}^b_{C,A} x + f_x \in \mathcal{M}$. We conclude that

$$\mathcal{M} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M}) = \{\widehat{\mathcal{O}}^b_{C,A} x + f_x \colon x \in \mathcal{D}(\Gamma)\}$$

and we can take the parametrizing map $X$ to be

$$X = \widehat{\mathcal{O}}^b_{C,A} + J(\widehat{\mathcal{C}}^b_{Z,B})^* (\mathcal{G}^J_{Z,B})^{-1} \Gamma \colon \mathcal{D}(\Gamma) \to \mathcal{M} \ominus (H^2_{\mathcal{U}} \cap \mathcal{M}) \subset L^2_{\mathcal{U}}. \tag{7.6}$$

We now apply the general principle above with $\mathcal{N} = \mathcal{M} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$ and $X$ as in (7.6). We therefore set

$$\begin{aligned} G^J_X := X^* J X &= \left((\widehat{\mathcal{O}}^b_{C,A})^* + \Gamma^* (\mathcal{G}^J_{Z,B})^{-1} \widehat{\mathcal{C}}^b_{Z,B} J\right) J \left(\widehat{\mathcal{O}}^b_{C,A} + J(\widehat{\mathcal{C}}^b_{Z,B})^* (\mathcal{G}^J_{Z,B})^{-1} \Gamma\right) \\ &= \mathcal{G}^J_{C,A} + \Gamma^* (\mathcal{G}^J_{Z,B})^{-1} \Gamma \end{aligned} \tag{7.7}$$

and the $J$-orthogonal projection onto $\mathcal{M} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})$ is given by

$$\begin{aligned} P_{\mathcal{M} \ominus_J (H^2_{\mathcal{U}} \cap \mathcal{M})} = &\left(\widehat{\mathcal{O}}^b_{C,A} + J(\widehat{\mathcal{C}}_{Z,B})^* (\mathcal{G}^J_{Z,B})^{-1} \Gamma\right) \cdot (G_X)^{-1} \\ &\cdot \left((\widehat{\mathcal{O}}^b_{C,A})^* J + \Gamma^* (\mathcal{G}^J_{Z,B})^{-1} \Gamma\right). \end{aligned} \tag{7.8}$$

Let us set $P_M = M_\Theta P_{H^2_{\mathcal{U}}} M_{\Theta^{-1}}$ equal to the $J$-orthogonal projection of $L^2$ onto $\mathcal{M}$. Therefore the operator $Q$ given by (7.3) can be computed as

$$\begin{aligned} Q &= J P_{H^2} - J P_{\mathcal{M}} \\ &= J(P_{H^2 \ominus_J \mathcal{M} \cap H^2_{\mathcal{U}}} + P_{\mathcal{M} \cap H^2_{\mathcal{U}}}) - J(P_{\mathcal{M} \ominus_J (\mathcal{M} \cap H^2_{\mathcal{U}})} + P_{\mathcal{M} \cap H^2_{\mathcal{U}}}) \\ &= J P_{H^2 \ominus_J (\mathcal{M} \cap H^2_{\mathcal{U}})} - J P_{\mathcal{M} \ominus_J (\mathcal{M} \cap H^2_{\mathcal{U}})} \\ &= \widehat{\mathcal{O}}^f_{B^*,Z^*} (\mathcal{G}^J_{Z,B})^{-1} (\widehat{\mathcal{O}}^f_{B^*,Z^*})^* - J \left(\widehat{\mathcal{O}}^b_{C,A} + J(\widehat{\mathcal{C}}^b_{Z,B})^* (\mathcal{G}^J_{Z,B})^{-1} \Gamma\right) \cdot \\ &\qquad (\mathcal{G}_X)^{-1} \cdot \left((\widehat{\mathcal{O}}^b_{C,A})^* + \Gamma^* (\mathcal{G}^J_{Z,B})^{-1} \widehat{\mathcal{C}}^b_{Z,B} J\right) J. \end{aligned} \tag{7.9}$$

Write $Q$ as a $2 \times 2$-block operator matrix with respect to the decomposition $L^2_{\mathcal{U}} = H^2_{\mathcal{U}} \oplus H^{2\perp}_{\mathcal{U}}$; the result is

$$Q = \begin{bmatrix} Q_{11} & -(\widehat{\mathcal{C}}^b_{Z,B})^*(\mathcal{G}^J_{Z,B})^{-1}\Gamma(\mathcal{G}^J_X)^{-1}(\widehat{\mathcal{O}}^b_{C,A})^* J \\ -J\widehat{\mathcal{O}}^b_{C,A}(\mathcal{G}^J_X)^{-1}\Gamma^*(\mathcal{G}^J_{Z,B})^{-1}\widehat{\mathcal{C}}^b_{Z,B} & -J\widehat{\mathcal{O}}^b_{C,A}(\mathcal{G}^J_X)^{-1}(\widehat{\mathcal{O}}^b_{C,A})^* J \end{bmatrix}$$

where we have set

$$Q_{11} = (\widehat{\mathcal{C}}^b_{Z,B})^*(\mathcal{G}^J_{Z,B})^{-1}\widehat{\mathcal{C}}^b_{Z,B} - (\widehat{\mathcal{C}}^b_{Z,B})^*(\mathcal{G}^J_{Z,B})^{-1}\Gamma(\mathcal{G}^J_X)^{-1}\Gamma^*(\mathcal{G}^J_{Z,B})^{-1}\widehat{\mathcal{C}}^b_{Z,B}.$$

We note the factorization

$$Q = M \begin{bmatrix} \mathcal{G}^J_{Z,B} - \Gamma(\mathcal{G}^J_X)^{-1}\Gamma^* & -\Gamma \\ -\Gamma^* & -\mathcal{G}^J_X \end{bmatrix} M^*$$

where we have set

$$M = \begin{bmatrix} (\widehat{\mathcal{C}}^b_{Z,B})^*(\mathcal{G}^J_{Z,B})^{-1} & 0 \\ 0 & J\widehat{\mathcal{O}}^b_{C,A}(\mathcal{G}^J_X)^{-1} \end{bmatrix}.$$

Since $M$ is injective, we see that positivity of $Q$ is equivalent to positivity of the middle factor

$$Q' = \begin{bmatrix} \mathcal{G}^J_{Z,B} - \Gamma(\mathcal{G}^J_X)^{-1}\Gamma^* & -\Gamma \\ -\Gamma^* & -\mathcal{G}^J_X \end{bmatrix}.$$

By a Schur complement analysis, positivity of $Q'$ is equivalent to

$$-\mathcal{G}^J_X > 0 \text{ and } (\mathcal{G}^J_{Z,B} - \Gamma(\mathcal{G}^J_X)^{-1}\Gamma^*) - \big((-\Gamma)(-\mathcal{G}^J_X)^{-1}(-\Gamma^*)\big) = \mathcal{G}^J_{Z,B} > 0. \tag{7.10}$$

On the other hand, $\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}$ given by (6.3) can be written as

$$\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}} = \begin{bmatrix} -\mathcal{G}^J_{C,A} & \Gamma^* \\ \Gamma & \mathcal{G}^J_{Z,B} \end{bmatrix}$$

where we are assuming that $\mathcal{G}^J_{Z,B}$ (as well as $\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}$) is invertible. A simple Schur-complement analysis with pivot equal to the $(2,2)$-block entry tells us that the positive-definiteness of $\Gamma_{\mathfrak{S},\mathfrak{S}^\times}$ is equivalent to

$$\mathcal{G}^J_{Z,B} > 0 \text{ and } -\mathcal{G}^J_{C,A} - \Gamma^*(\mathcal{G}^J_{Z,B})^{-1}\Gamma > 0. \tag{7.11}$$

Recalling now the definition (7.7) of $\mathcal{G}^J_X$ we see that (7.11) and (7.10) are equivalent. This finally gives us the equivalence of (2) and (3) for the case that $\mathcal{G}^J_{Z,B}$ is invertible.

Finally we note that $\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}$ being positive-semidefinite implies that $\Gamma$ is bounded. To see this, apply the Cauchy-Schwarz inequality for the positive-semidefinite form $\langle \cdot, \cdot \rangle_{\Gamma_{\mathfrak{S},\mathfrak{S}^\times}}$ on $\mathcal{D}(\Gamma) \oplus \mathcal{D}(\Gamma^*)$ given by

$$\left\langle \begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} f' \\ g' \end{bmatrix} \right\rangle_{\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}} = \left\langle \begin{bmatrix} -\mathcal{G}^J_{C,A} & \Gamma^* \\ \Gamma & \mathcal{G}^J_{Z,B} \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}, \begin{bmatrix} f' \\ g' \end{bmatrix} \right\rangle_{\mathcal{X}_{\mathcal{P}} \oplus \mathcal{X}_{\mathcal{Z}}}$$

to conclude that

$$\left| \left\langle \begin{bmatrix} f \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ g' \end{bmatrix} \right\rangle_{\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}} \right| \le \left\langle \begin{bmatrix} f \\ 0 \end{bmatrix}, \begin{bmatrix} f \\ 0 \end{bmatrix} \right\rangle^{1/2}_{\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}} \cdot \left\langle \begin{bmatrix} 0 \\ g' \end{bmatrix}, \begin{bmatrix} 0 \\ g' \end{bmatrix} \right\rangle^{1/2}_{\Gamma_{\mathfrak{S},\mathfrak{S}^{\perp J}}}$$

or

$$|\langle \Gamma f, g' \rangle_{\mathcal{X}_{\mathcal{Z}}}| \le \|(-\mathcal{G}^J_{C,A})^{1/2} f\| \cdot \|(\mathcal{G}^J_{Z,B})^{1/2} g'\| \le \|(-\mathcal{G}^J_{C,A})^{1/2}\| \|(\mathcal{G}^J_{Z,B})^{1/2}\| \|f\| \|g'\|$$

from which we conclude that

$$\|\Gamma\| \le \|(-\mathcal{G}^J_{C,A})^{1/2}\| \|(\mathcal{G}^J_{Z,B})^{1/2}\| < \infty.$$

This concludes the proof of Theorem 7.1. □

## 8. Generalized Schur-Nevanlinna-Pick interpolation

In this section we still assume that $\dim \mathcal{U} < \infty$ but in addition we assume that $\mathcal{U}$ has a decomposition as

$$\mathcal{U} = \begin{bmatrix} \mathcal{U}_o \\ \mathcal{U}_i \end{bmatrix}$$

The notation is to suggest that $\mathcal{U}_i$ is an *input* space and $\mathcal{U}_o$ is an *output* space and we are interested in studying the space $H^\infty_{\mathcal{L}(\mathcal{U}_i,\mathcal{U}_o)}$ of $\mathcal{L}(\mathcal{U}_i,\mathcal{U}_o)$-valued analytic functions on the unit disk $\mathbb{D}$. We are particularly interested in the closed unit ball of this space, often called the *Schur class*, which we shall denote by

$$\mathcal{S}(\mathcal{U}_i,\mathcal{U}_o) = \{s \in H^\infty_{\mathcal{L}(\mathcal{U}_i,\mathcal{U}_o)} : \|s\|_\infty \le 1\}.$$

We follow the usual conventions and abbreviate $\mathcal{S}(\mathcal{U}_i,\mathcal{U}_i)$ to $\mathcal{S}(\mathcal{U}_i)$ and similarly with $\mathcal{S}(\mathcal{U}_o)$. We assume that we are given a two-sided inner function $b_o \in \mathcal{S}_{\mathcal{L}(\mathcal{U}_o)}$, a two-sided inner function $b_i \in \mathcal{S}(\mathcal{U}_i)$, and a function $s_0 \in H^\infty_{\mathcal{L}(\mathcal{U}_i,\mathcal{U}_o)}$. The *generalized Schur interpolation problem* (with data set $b_o, b_i, s_0$) GSIP$(b_o, b_i, s_0)$ then is: *find* $s \in \mathcal{S}(\mathcal{U}_i,\mathcal{U}_o)$ *so that* $b_o^{-1}(s - s_0)b_i^{-1} \in H^\infty_{\mathcal{L}(\mathcal{U}_i,\mathcal{U}_o)}$. We introduce a notation for the set of all solutions:

$$\mathcal{S}(b_o, b_i, s_0) = \{s \in \mathcal{S}(\mathcal{U}_i,\mathcal{U}_o) : s \text{ solves GSIP}(b_o, b_i, s_0)\}.$$

There is an equivalent state-space formulation of the problem GSIP$(b_o, b_i, s_0)$, called a bitangential interpolation problem (BTIP), which we now discuss. Let $A \colon \mathcal{X}_R \to \mathcal{X}_R$, $U \colon \mathcal{X}_R \to \mathcal{U}_i$, $V \colon \mathcal{X}_R \to \mathcal{U}_o$, $Z \colon \mathcal{X}_L \to \mathcal{X}_L$, $X : \mathcal{U}_o \to \mathcal{X}_L$, $Y : \mathcal{U}_i \to \mathcal{X}_L$ and $\Gamma \colon \mathcal{X}_R \to \mathcal{X}_L$ be bounded linear operators. We say that

$$\mathcal{I} = (U, V, A, Z, X, Y, \Gamma) \tag{8.1}$$

is an *admissible interpolation data set* if $(U, A)$ is exactly observable, $(Z, X)$ is exactly controllable, $A$ and $Z$ are strongly bi-stable, and $\Gamma$ satisfies the Sylvester equation $\Gamma A - Z\Gamma = XV - YU$. It is immediate that if $(U, A)$ is exactly observable then $([\begin{smallmatrix} V \\ U \end{smallmatrix}], A)$ is exactly observable, and if $(Z, X)$ is exactly controllable then $(Z, [\begin{smallmatrix} X & -Y \end{smallmatrix}])$ is exactly controllable. Hence

$$\mathfrak{S}_{\mathcal{I}} = \left( \begin{bmatrix} V \\ U \end{bmatrix}, A, Z, \begin{bmatrix} X & -Y \end{bmatrix}, \Gamma \right) \tag{8.2}$$

is an associated admissible Sylvester data set. The *bitangential interpolation problem* (with data set $\mathcal{I}$) BTIP($\mathcal{I}$) then is: *Find all $s \in \mathcal{S}(\mathcal{U}_i, \mathcal{U}_o)$ so that*

$$P_{H^{2\perp}_{\mathcal{U}_o}} M_s \widehat{\mathcal{O}}^b_{U,A} = \widehat{\mathcal{O}}^b_{V,A}, \qquad \widehat{\mathcal{C}}^b_{Z,X} M_s|_{H^2_{\mathcal{U}_i}} = \widehat{\mathcal{C}}^b_{Z,Y}, \qquad \widehat{\mathcal{C}}^b_{Z,X} P_{H^2_{\mathcal{U}_o}} M_s \widehat{\mathcal{O}}^b_{U,A} = \Gamma.$$

We introduce a notation for the set of all solutions:

$$\mathcal{S}(\mathcal{I}) = \{ s \in \mathcal{S}(\mathcal{U}_i, \mathcal{U}_o) \colon s \text{ solves BTIP}(\mathcal{I}) \}.$$

Given $(b_o, b_i, s_0)$ for the GSIP, we may construct $\mathcal{I}$ as follows. *Let $\mathcal{M}$ be the full-range $M_z$-simply-invariant subspace*

$$\mathcal{M}_{b_o, b_i, s_0} = \begin{bmatrix} b_o & s_0 b_i^{-1} \\ 0 & b_i^{-1} \end{bmatrix} H^2_{\mathcal{U}} \tag{8.3}$$

*and construct an admissible Sylvester data set $\mathfrak{S} = (C, A, Z, B, \Gamma)$ so that*

$$\mathcal{M}_{b_o, b_i, s_0} = \mathcal{M}_{\mathfrak{S}}.$$

*Then $\mathfrak{S}$ has the form*

$$\mathfrak{S} = \left( \begin{bmatrix} U \\ V \end{bmatrix}, \quad A, \quad Z, \quad \begin{bmatrix} X & -Y \end{bmatrix}, \quad \Gamma \right)$$

*where $\mathcal{I} = (U, V, A, Z, X, Y, \Gamma)$ is an admissible interpolation data set. Moreover, the solution sets for the two problems are the same:* BTIP*($\mathcal{I}$)* = GSIP*($b_o, b_i, s_0$). Conversely, given an admissible interpolation data set $\mathcal{I}$ for a* BTIP*, it is possible to construct a data set $(b_o, b_i, s_0)$ for a* GSIP *so that $\mathcal{S}(b_0, b_i, s_0) = \mathcal{S}(\mathcal{I})$.*

Parametrizing solutions of such problems is closely related to $J$-inner functions where we take $J = \begin{bmatrix} I_{\mathcal{U}_0} & 0 \\ 0 & -I_{\mathcal{U}_i} \end{bmatrix}$. The following result can be proved by combining the results of this paper with the approach to interpolation from [20, 19, 23].

**Theorem 8.1.** *Given the data $(b_o, b_i, s_0)$ for a* GSIP *(respectively, the data $\mathcal{I}$ as in* (8.1) *for a* BTIP*), form the subspace $\mathcal{M} := \mathcal{M}_{b_o, b_i, s_0} = \mathcal{M}_{\mathfrak{S}_{b_o, b_i, s_0}}$ as in* (8.3) *(respectively, the subspace $\mathcal{M} := \mathcal{M}_{\mathcal{I}} = \mathcal{M}_{\mathfrak{S}_{\mathcal{I}}}$ where $\mathfrak{S}_{\mathcal{I}}$ is the admissible Sylvester set associated with $\mathcal{I}$ as in* (8.2)*). Then the interpolation problem* GSIP *(respectively* BTIP*) has solutions if and only if the coupling matrix $\Gamma := \Gamma_{\mathfrak{S}_{b_o, b_i, s_0}, (\mathfrak{S}_{b_o, b_i, s_0})^{\perp J}}$ (respectively, $\Gamma := \Gamma_{\mathfrak{S}_{\mathcal{I}}, (\mathfrak{S}_{\mathcal{I}})^{\perp J}}$) is positive semidefinite. In case $\Gamma$ is strictly positive definite, let*

$$\Theta(z) = \begin{bmatrix} \Theta_{11}(z) & \Theta_{12}(z) \\ \Theta_{21}(z) & \Theta_{22}(z) \end{bmatrix}$$

*be a $J$-inner function such that $L^2$-clos $\Theta H^\infty_{\mathcal{U}} = \mathcal{M}$. Then a function $S \in \mathcal{S}(\mathcal{U}_i, \mathcal{U}_o)$ solves the interpolation problem* GSIP *(or* BTIP*) if and only if $S$ has the form*

$$s(z) = (\Theta_{11}(z)g(z) + \Theta_{12}(z))(\Theta_{21}(z)g(z) + \Theta_{22}(z))^{-1} \tag{8.4}$$

*for some free-parameter Schur-class function $g \in \mathcal{S}(\mathcal{U}_i, \mathcal{U}_o)$.*

The parametrization result in Theorem 8.1 handles only the case where the coupling operator (which can be identified with the *Pick matrix* in the classical case) is (strictly) positive-definite. There is a more general situation where the coupling operator is positive-semidefinite but still has no kernel. In this case the associated shift-invariant subspace $\mathcal{M}$ contains no isotropic subspace, i.e., $\mathcal{M} \cap \mathcal{M}^{\perp J} = \{0\}$, but $\mathcal{M} \dot{+} \mathcal{M}^{\perp J}$ is only dense in $L^2_{\mathcal{U}}$ rather than being equal to all of $L^2_{\mathcal{U}}$. This corresponds to the so-called *completely indeterminate* case of the interpolation problem (see e.g. [4]) where *there exists a point $\omega \in \mathbb{D}$ where both $b_o(\omega)$ and $b_i(\omega)$ are invertible such that, for every nonzero vector $\eta \in \mathcal{U}_i$, there exists a solution $s \in \mathcal{S}(b_o, b_i, s_0)$ so that $s(\omega)\eta \neq s_0(\omega)\eta$*. In this case there is a still a linear-fractional parametrization of the set of all solutions $\mathcal{S}(b_o, b_i, s_0)$ via a $J$-inner function $\Theta$ as in Theorem 8.1, but the so-called *resolvent matrix* $\Theta$ in general is not $L^2$-regular. The precise class of $J$-inner functions arising in this way is now called the class of *Arov-regular* $J$-inner functions and has a number of characterizations (see [2, 3]). The subclass of *strongly Arov-regular* $J$-inner functions has come up in a number of applications and has a number of characterizations (see [4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15]), one of which is that the associated set of solutions of the associated interpolation problem $\mathcal{S}(b_o, b_i, s_0)$ includes at least one interpolant $s$ with $\|s\|_\infty < 1$ (the *strictly indeterminate case*—see [4]). Our contribution here is to draw attention to another characterization via the $L^2$-regular property which we have developed in this paper.

**Theorem 8.2.** *Let $\Theta$ be a $J$-inner function on $\mathbb{D}$. Then the following are equivalent.*

1. *$\Theta$ is strongly Arov-regular, i.e., the set of functions $s$ of the form* (8.4) *for some $g \in \mathcal{S}(\mathcal{U}_i, \mathcal{U}_o)$ includes a function $s \in \mathcal{S}(\mathcal{U}_i, \mathcal{U}_o)$ with $\|s\|_\infty < 1$.*
2. *The reproducing kernel Hilbert space $\mathcal{H}(K_\Theta)$ associated with the kernel $K_\Theta$* (7.2) *is contained in $L^2_{\mathcal{U}}$:*
$$\mathcal{H}(K_\Theta) \subset L^2_{\mathcal{U}}.$$
3. *$\Theta \in L^2_{\mathcal{U}}$ and the function*
$$\Delta(z) := (\Theta_{21}(z) + \Theta_{22}(z))^*(\Theta_{21}(z) + \Theta_{22}(z))$$
*satisfies the Muckenhoupt condition: if, for $I$ an arc on the unit circle, we set $|I|$ equal to the length of the arc $I$ and $A_I(\Delta) = \frac{1}{|I|}\int_I \Delta(z)\,|dz|$ equal to the average of $\Delta(z)$ over $I$, then*
$$\sup_I \left| A_I(\Delta)^{1/2} A_I(\Delta^{-1})^{1/2} \right| < \infty.$$
4. *$\Theta$ is $L^2$-regular, i.e., $\Theta \in L^2_{\mathcal{L}(\mathcal{U})}$ and $M_\Theta P_{H^2_{\mathcal{U}}} M_{\Theta^{-1}} = M_\Theta P_{H^2_{\mathcal{U}}} J M_{\Theta^*} J$ extends to a bounded operator on $L^2_{\mathcal{U}}$.*

*Proof.* The equivalence (1) $\Longleftrightarrow$ (2) is established in [4] while the equivalence (1) $\Longleftrightarrow$ (3) is carried out in [9]. We shall show that (4) $\Longrightarrow$ (2) and that (1) $\Longrightarrow$ (4).

**(4) $\Longrightarrow$ (2):** By assumption $P_{\mathcal{M}} = M_\Theta P_{H^2_{\mathcal{U}}} M_{\Theta^{-1}}$ is bounded, so $Q$ defined by (7.3) is also bounded. We know from Theorem 7.1 that $Q$ is positive-semidefinite since $\Theta$ is $J$-inner. In more detail, from the proof of this fact coming out of the identity (7.4), we read off that the reproducing kernel Hilbert space $\mathcal{H}(K_\Theta)$ can be characterized as Ran $Q^{1/2}$ with lifted norm

$$\|Q^{1/2} f\|_{\mathcal{H}(K_\Theta)} = \|P_{(\mathrm{Ker}\ Q)^\perp} f\|_{L^2_{\mathcal{U}}}.$$

In particular, it follows that $\mathcal{H}(K_\Theta) \subset L^2_{\mathcal{U}}$.

**(1) $\Longrightarrow$ (4):** By the Grassmannian approach to interpolation from [20], we know that a strictly contractive solution of GSIP$(b_o, b_i, s)$ is equivalent to the subspace $\mathcal{M}$ containing a subspace $\mathcal{G}$ which is (1) uniformly negative, (2) maximal negative as a subspace of $\mathcal{M}$, and (3) shift-invariant. Just the first two properties together imply that $\mathcal{M}$ is orthocomplemented in $(L^2_{\mathcal{U}}, J)$ (see [27, Lemma V.7.5]). This in turn is equivalent to the existence of a bounded $J$-orthogonal projection $P_{\mathcal{M}}$. Necessarily $P_{\mathcal{M}}$ is an extension of $M_\Theta P_{H^2_{\mathcal{U}}} M_{\Theta^{-1}}$, i.e., $\Theta$ is $L^2$-regular. □

## References


[1] D. Alpay and I. Gohberg, Unitary rational matrix functions, in *Topics in Interpolation Theory and Rational Matrix-Valued Functions* (Ed. I. Gohberg), pp. 175-222, **OT 33**, Birkhäuser-Verlag, Basel-Boston, 1988.

[2] D.Z. Arov, $\gamma$-generating matrices, $j$-inner matrix-functions and related extrapolation problems, *Teor. Funktsii Funktsional. Anal. i Prilozhen.*, I, **51** (1989), 61-67; II, **52** (1989), 103-109; III, **53** (1990), 57-64; English transl., *J. Soviet Math.* I, **52** (1990), 3487-3491; II, **52** (1990), 3421-3425; III, **58** (1992), 532, 537.

[3] D.Z. Arov, $\gamma$-generating matrices, $j$-inner matrix-functions and related extrapolation problems, IV, *Mathematical Physics, Analysis, Geometry* **2** (1995), 3-14.

[4] D.Z. Arov and H. Dym, $J$-inner matrix functions, interpolation and inverse problems for canonical systems, I: Foundations, *Integral Equations and Operator Theory* **29** (1997), 373-454.

[5] D.Z. Arov and H. Dym, $J$-inner matrix functions, interpolation and inverse problems, II: the inverse monodromy problem, *Integral Equations and Operator Theory* **36** (2000), 11-70.

[6] D.Z. Arov and H. Dym, $J$-inner matrix functions, interpolation and inverse problems for canonical systems, III: More on the inverse monodromy problem, *Integral Equations and Operator Theory* **36** (2000), 127-181.

[7] D.Z. Arov and H. Dym, $J$-inner matrix functions, interpolation and inverse problems for canonical systems, IV: Direct and inverse bitangential input scattering problem, *Integral Equations and Operator Theory* **43** (2002), 1–67.

[8] D.Z. Arov and H. Dym, $J$-inner matrix functions, interpolation and inverse problems for canonical systems, V: The inverse input scattering problem for Wiener class and

rational $p \times q$ input scattering matrices, *Integral Equations and Operator Theory* **43** 2002, 68-129.

[9] D.Z. Arov and H. Dym, Matricial Nehari problems, $J$-inner matrix functions and the Muckenhoupt condition, *J. Functional Analysis* **280** (2003), 387-399.

[10] D.Z. Arov and H. Dym, Matricial Nehari problems, $J$-inner matrix functions and $\gamma$-generating matrices, *J. Math. Anal. Appl.* **280** (2003), 387-399.

[11] D.Z. Arov and H. Dym , Strongly regular $J$-inner matrix functions and related problems, in *Current Trends in Operator Theory and its Applications* (Ed. J.A. Ball, J.W. Helton, J. Klaus and L. Rodman), pp.79-106, **OT 149** Birkhäuser-Verlag, Basel-Boston, 2004.

[12] D.Z. Arov and H. Dym, Strongly regular $J$-inner matrix valued functions and inverse problems for canonical systems, *Recent Advances in Operator Theory and its Applications* (Ed. M.A. Kaashoek, S. Seatzu and C. van der Mee), pp. 101-160, **OT 160** Birkhäuser-Verlag, Basel-Boston, 2005.

[13] D.Z. Arov and O.J. Staffans, State/signal linear time-invariant systems theory, Part I: Discrete time systems, in *The State Space Method Generalizations and Applications* (Ed. D. Alpay and I. Gohberg), pp. 115-177, OT **161**, Birkhäuser-Verlag, Basel-Boston-Berlin, 2005.

[14] Z.D. Arova, $J$-unitary nodes with strongly regular $J$-inner characteristic matrix functions, *Methods in Functional Analysis and Topology* **6** (2000), no. 3, 9-23.

[15] Z.D. Arova, On Livšic-Brodskii nodes with strongly regular $J$-inner characteristic matrix functions, in *Recent Advances in Operator Theory and Related Topics* (Ed. L. Kérchy, C. Foias, I. Gohberg and H. Langer) pp. 83-96, **OT 127** Birkhäuser-Verlag, Basel-Boston, 2001.

[16] Z.D. Arova, *Operator Nodes with Strongly Regular Characteristic Function*, Academisch Proefschrift, Vrije Universiteit, Amsterdam, 2003.

[17] J.A. Ball, Nevanlinna-Pick interpolation: generalizations and applications, in *Recent Results in Operator Theory, Vol. I* (Ed. J.B. Conway and B.B. Morrel), pp. 51-94, Longman Sci. Tech., Harlow, 1988.

[18] J.A. Ball, N. Cohen and A.C.M. Ran, Inverse spectral problems for regular improper rational matrix functions, in *Topics in Interpolation Theory and Rational Matrix-Valued Functions* (Ed. I. Gohberg), pp. 123-173, **OT 33** Birkhäuser-Verlag, Basel-Boston, 1988.

[19] J.A. Ball, I. Gohberg and L. Rodman, *Interpolation of Rational Matrix Functions*, **OT 45**, Birkhäuser-Verlag, Basel-Boston, 1990.

[20] J.A. Ball and J.W. Helton, A Beurling-Lax theorem for the Lie group $U(m,n)$ which contains most classical interpolation, *J. Operator Theory* **9** (1983), 107-142.

[21] J.A. Ball and J.W. Helton, Beurling-Lax representations using classical Lie groups with many applications II: $GL(n,\mathbb{C})$ and Wiener-Hopf factorization, *Integral Equations and Operator Theory* **7** no. 3 (1984), 291-309.

[22] J.A. Ball and J.W. Helton, Shift invariant subspaces, passivity, reproducing kernels and $H^\infty$-optimization, in *Contributions to Operator Theory and its Applications* (Ed. J.W. Helton and L. Rodman), pp. 265-310, **OT 35** Birkhäuser-Verlag, Basel-Boston, 1988.

[23] J.A. Ball, K.M. Mikkola and A.J. Sasane, State-space formulas for the Nehari-Takagi problem for nonexponentially stable infinite-dimensional systems, *SIAM J. Control Optim.* **44** No. 2 (2006), 531-563.

[24] J.A. Ball and A.C.M. Ran, Global inverse spectral problems for rational matrix functions, *Linear Algebra and its Applications* **86** (1987), 237-282.

[25] J.A. Ball and A.C.M. Ran, Local inverse spectral problems for rational matrix functions, *Integral Equations and Operator Theory* **10** (1987), 349-415.

[26] A. Beurling, On two problems concerning linear transformations in Hilbert space, *Acta Math.* **81** (1949), 239–255.

[27] J. Bognár, *Indefinite Inner Product Spaces*, Springer-Verlag, New Your-Heidelberg-Berlin, 1974.

[28] R.G. Douglas and J.W. Helton, Inner dilations of analytic matrix functions and Darlington synthesis, *Acta Sci. Math. (Szeged)* **34** (1973), 61-67.

[29] R.G. Douglas, H.S. Shapiro and A.L. Shields, Cyclic vectors of the backward shift, *Ann. Inst. Fourier (Grenoble)*, **20** (1970), 37-76.

[30] H. Dym, *J-Contractive Matrix Functions, Reproducing Kernel Hilbert Spaces and Interpolation*, CBMS Regional Conference Series **71**, American Mathematical Society, Providence, 1989.

[31] C. Foias and A.E. Frazho, *The Commutant Lifting Approach to Interpolation Problems*, **OT 44** Birkhäuser-Verlag, Basel-Boston, 1990.

[32] C. Foias, A.E. Frazho, I. Gohberg and M.A. Kaashoek, *Metric Constrained Interpolation, Commutant Lifting and Systems*, **OT 100** Birkhäuser-Verlag, Basel-Boston, 1998.

[33] P.R. Halmos, Shifts on Hilbert spaces, *J. Reine Angew. Math.* **208** (1961), 102–112.

[34] J.W. Helton, J.A. Ball, C.B. Johnson and J.N. Palmer, *Operator Theory, Analytic Functions, Matrices and Electrical Engineering*, CBMS Regional Conference Series, **68**, American Mathematical Society, Providence, 1987.

[35] O. Iftume, M.A. Kaashoek and A. Sasane, A Grassmannian band method approach to the Nehari-Takagi problem, *J. Math. Anal. Appl.* **310** (2005), 97-115.

[36] P.D. Lax, Translation invariant spaces, *Acta Math.* **101** (1959), 163–178.

[37] J. Malinen, O.J. Staffans and G. Weiss, When is a linear system conservative? *Quarterly of Applied Mathematics* **LXIV** no. 1 (2006), 61-91.

[38] B. Sz.-Nagy and C. Foias, *Harmonic Analysis of Operators on Hilbert Space*, North Holland Publishing Company, Amsterdam-London, 1970.

[39] A.A. Nudelman, On a new problem of moment type, *Soviet Math. Doklady* **18** (1977), 507-510.

[40] M. Rosenblum and J. Rovnyak, *Hardy Classes and Operator Theory*, Oxford University Press, New York, 1985.

[41] W. Rudin, *Functional Analysis, Second Edition*, International Series in Pure and Applied Mathematics, McGraw Hill, Boston, 1991.

[42] O.J. Staffans, *Well-Posed Linear Systems*, Cambridge University Press, Cambridge, UK, 2004.

Joseph A. Ball
Department of Mathematics
Virginia Tech
Blacksburg, Virginia 24061
e-mail: ball@math.vt.edu

Michael W. Raney
Department of Mathematics
Virginia Tech
Blacksburg, Virginia 24061
e-mail: michaelraney@verizon.net

Communicated by Daniel Alpay
Submitted: August 20, 2006
Accepted: September 13, 2006