BEHAVIORAL REALIZATIONS USING COMPANION MATRICES AND THE SMITH FORM*

TIMOTHY H. HUGHES†

Abstract. Classical procedures for the realization of transfer functions are unable to represent uncontrollable behaviors. In this paper, we use companion matrices and the Smith form to derive explicit observable realizations for a general (not necessarily controllable) linear time-invariant behavior. We then exploit the properties of companion matrices to efficiently compute trajectories, and the solutions to Lyapunov equations, for the realizations obtained. The results are motivated by the important role played by uncontrollable behaviors in the context of physical systems such as passive electrical and mechanical networks [4, 11, 12, 10].

Key words. Behavioral theory, Controllability, Realization, Linear systems, Algebraic systems theory

AMS subject classifications. 34H05, 93B05, 93B20, 93B25, 93C05, 93C35

1. Introduction. A natural way to describe the behavior of physical systems is with a set of relationships between the system's variables. For linear time-invariant systems, these relationships take the general form $R(\frac{d}{dt})\mathbf{w} = 0$, where R is some real polynomial matrix which corresponds to the laws of the system, and the solutions $\mathbf{w}$ to $R(\frac{d}{dt})\mathbf{w} = 0$ correspond to those evolutions in time of the system's variables which are permitted by these laws. In many cases, the system's variables are partitioned into *inputs* $\mathbf{u}$ and *outputs* $\mathbf{y}$ which satisfy a relationship of the form $R_1(\frac{d}{dt})\mathbf{y} = R_2(\frac{d}{dt})\mathbf{u}$ where R_1 and R_2 are real polynomial matrices and R_1 is non-singular. In fact, any linear time-invariant behavior can be represented in this form [17, Theorem 2.5.23 and Corollary 3.3.23].

On the other hand, in optimization and control, it is common for the analysis to proceed from relationships of the form $\frac{d\mathbf{x}}{dt} = A\mathbf{x} + B\mathbf{u}$, and $\mathbf{y} = C\mathbf{x} + D(\frac{d}{dt})\mathbf{u}$, where A, B, and C are real matrices, and D is a real polynomial matrix (which may also be a real matrix). Indeed, the solutions to many fundamental problems in optimization and control use such representations, e.g. the $\mathscr{H}_2$ and $\mathscr{H}_\infty$ optimal control problems [5]. There is also a significant advantage to such a representation insofar as simulation is concerned: given a (sufficiently smooth) $\mathbf{u}$ and a real $\mathbf{x}(0)$, there is a unique $\mathbf{x}$ which satisfies $\frac{d\mathbf{x}}{dt} = A\mathbf{x} + B\mathbf{u}$ (this can be computed with the variation of the constants formula [17, Section 4.5]), whereupon we obtain a unique solution for $\mathbf{y}$.

This motivates the *behavioral realization* problem: given polynomial matrices R_1 and R_2, find real matrices A, B, C, and a real polynomial matrix D, such that the solutions to $R_1(\frac{d}{dt})\mathbf{y} = R_2(\frac{d}{dt})\mathbf{u}$ are given by $\mathbf{y} = C\mathbf{x} + D(\frac{d}{dt})\mathbf{u}$ with $\frac{d\mathbf{x}}{dt} = A\mathbf{x} + B\mathbf{u}$. There is a crucial distinction between this problem and classical realization procedures which are typically focussed on realizing the transfer function $G = R_1^{-1}R_2$ (e.g. [13, 9, 22]). These classical procedures are unable to realize uncontrollable behaviors, for example the driving-point behavior of the network in Fig. 1. From [12, Section 7], this is the set of solutions to $g(\frac{d}{dt})\left(q(\frac{d}{dt})v - p(\frac{d}{dt})i\right) = 0$ for $g(\xi) = \xi + 1$, $p(\xi) = \xi^2 + \xi + 1$, and $q(\xi) = \xi^2 + \xi + 4$, and it is behaviorally uncontrollable (see Section 2). The

*This research was conducted while the author was the Henslow research fellow at Fitzwilliam College, University of Cambridge, U.K., supported by the Cambridge Philosophical Society, http://www.cambridgephilosophicalsociety.org

†Cambridge University Engineering Department, Trumpington Street, Cambridge, UK, CB2 1PZ. (thh22@cam.ac.uk).

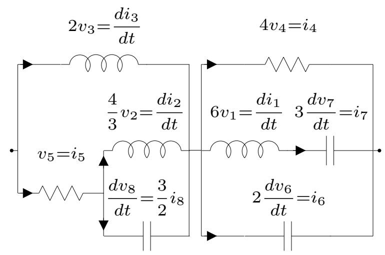

FIG. 1. *The network of Bott and Duffin for achieving the impedance* $(\xi^2 + \xi + 1)/(\xi^2 + \xi + 4)$

transfer function p/q is thus insufficient for determining the driving-point behavior of this network as it does not capture the polynomial g. We note that this network was found by Bott and Duffin in [3], and it contains the least possible number of energy storage elements (inductors and capacitors) among all series-parallel resistor, inductor, capacitor networks with impedance p/q (see [11]).

One objective of this paper is to derive realizations of (not necessarily controllable) behaviors which can be efficiently computed using algorithms available in symbolic algebra programs. A second objective is to construct realizations using companion matrices. These enable the efficient computation of matrix exponentials [15, 14], and the solutions to Lyapunov and Sylvester equations [6, 1]. In particular, we extend the results in [15, 14, 6, 1] to efficiently compute trajectories, and the solutions to Lyapunov equations, for the realizations in this paper. An example is provided in which we compute the observability and controllability gramians for a stable system.

The paper is structured as follows. We begin with some background on linear time-invariant differential behaviors in Section 2. In Section 3, we compare the results in this paper to past approaches to the realization of transfer functions and behaviors. In Section 4, we provide an explicit realization for the behavior defined by $R_1(\frac{d}{dt})\mathbf{y} = R_2(\frac{d}{dt})\mathbf{u}$. The properties of this realization are examined in Section 5. In particular, we show that it is observable, and that it is controllable if and only if it is representing a controllable behavior. We also provide a second realization with these same properties. Then, in Section 6, we extend results from [15, 14, 6, 1] on efficient computations with companion matrices, and we apply these results to the realizations in this paper. Finally, in Section 7, we derive realizations for the behavior defined by $R(\frac{d}{dt})\mathbf{w} = 0$.

1.1. Notation. $\mathbb{R}$ (resp. $\mathbb{C}$) denotes the real (resp. complex) numbers, and $\mathbb{R}[\xi]$ (resp. $\mathbb{R}(\xi)$) the space of real polynomials (resp. real rational functions). We say $R \in \mathbb{R}(\xi)$ is proper (resp. strictly proper) if R is bounded (resp. zero) at infinity, and we denote the space of proper real rational functions by $\mathbb{R}_p(\xi)$. Let $\mathbb{F}$ be one of $\mathbb{R}, \mathbb{C}, \mathbb{R}[\xi], \mathbb{R}(\xi)$, or $\mathbb{R}_p(\xi)$. Then $\mathbb{F}^{m \times n}$ denotes the matrices with m rows and n columns whose entries are all from $\mathbb{F}$, and we write $\mathbb{F}^{\bullet \times \bullet}$ (resp. $\mathbb{F}^{\bullet}$) when these numbers are immaterial. I_m denotes the $m \times m$ identity matrix, 0_m a column vector of m zeros, $0_{m \times n}$ an $m \times n$ matrix of zeros, and the dimensions are occasionally omitted when clear from the context. Finally, $\operatorname{col}\begin{pmatrix} M_1 & \cdots & M_n \end{pmatrix} = \begin{bmatrix} M_1^T & \cdots & M_n^T \end{bmatrix}^T$, and $\operatorname{diag}\begin{pmatrix} M_1 & \cdots & M_n \end{pmatrix}$ denotes a block diagonal matrix with $M_1, \ldots, M_n \in \mathbb{F}^{\bullet \times \bullet}$ appearing in this order in the main diagonal blocks.

2. Linear time-invariant differential behaviors. In this paper, we consider behaviors defined as the sets of solutions of linear time-invariant differential equations, and we refer to elements from the behavior as trajectories. This is in keeping with the behavioral approach to mathematical systems theory [17]. Here, we summarise aspects of behavioral theory which are required in this paper.

Following [17, Section 2.3.2], we interpret differentiation in a weak sense, we allow solutions from the space of locally integrable functions, and we consider two functions to be identical if they are equal except on a set of measure zero. Such assumptions are typical in linear systems theory. Thus, a behavior has the general form:

$$\mathcal{B} = \left\{ \mathbf{w} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \,\middle|\, R(\tfrac{d}{dt})\mathbf{w} = 0,\ R \in \mathbb{R}^{l \times n}[\xi] \right\}. \tag{2.1}$$

On occasion, we consider the subspace of $\mathcal{B}$ comprising the infinitely differentiable solutions to $R(\frac{d}{dt})\mathbf{w} = 0$, which we denote $\mathcal{B} \cap \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^n)$. Note that any conventional (strong) solution to a differential equation is also a weak solution (see [17, Theorem 2.3.11]), so in our examples we will usually interpret differentiation conventionally.

It is often convenient (and always possible) to represent behaviors in the form:

$$\mathcal{B}_{i/o} = \left\{ \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \,\middle|\, \begin{array}{l} R_1(\frac{d}{dt})\mathbf{y} = R_2(\frac{d}{dt})\mathbf{u}, \\ R_1 \in \mathbb{R}^{m \times m}[\xi] \text{ and } R_2 \in \mathbb{R}^{(n-m) \times m}[\xi], \\ \text{and } R_1(\lambda) \text{ non-singular for almost all } \lambda \in \mathbb{C} \end{array} \right\}. \tag{2.2}$$

As will be shown in Section 7, for any $\mathcal{B}$ of the form (2.1), there exists an invertible matrix $T := \begin{bmatrix} T_1 & T_2 \end{bmatrix}$ such that $\mathbf{w} \in \mathcal{B}$ if and only if $\mathbf{w} = T_1\mathbf{y} + T_2\mathbf{u}$ where $\text{col}\begin{pmatrix} \mathbf{y} & \mathbf{u} \end{pmatrix} \in \mathcal{B}_{i/o}$. For the behavior $\mathcal{B}_{i/o}$, given $\mathbf{u} \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{n-m})$, there always exists a $\mathbf{y} \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^m)$ such that $\text{col}\begin{pmatrix} \mathbf{y} & \mathbf{u} \end{pmatrix} \in \mathcal{B}_{i/o} \cap \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^n)$. If, in addition, the transfer function $R_1^{-1}R_2$ is proper, then, given a $\mathbf{u} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^{n-m})$, there always exists a $\mathbf{y} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^m)$ such that $\text{col}\begin{pmatrix} \mathbf{y} & \mathbf{u} \end{pmatrix} \in \mathcal{B}_{i/o}$ [17, Section 3.3]. Nevertheless, in many physical systems, it is natural to consider non-proper transfer function, e.g. the transfer function from current to voltage for an inductor. Accordingly, we refer to $\mathbf{u}$ and $\mathbf{y}$ in (2.2) as an input and output, respectively, irrespective of whether $R_1^{-1}R_2$ is proper (this is in contrast with [17]).

In Section 4 of this paper, we will seek a realization of the behavior $\mathcal{B}_{i/o}$ of the form $\mathcal{B}_{i/o} = \{\text{col}\begin{pmatrix} \mathbf{y} & \mathbf{u} \end{pmatrix} \mid \exists \mathbf{x} \text{ with } \text{col}\begin{pmatrix} \mathbf{y} & \mathbf{u} & \mathbf{x} \end{pmatrix} \in \mathcal{B}_s\}$ for

$$\mathcal{B}_s = \left\{ \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{x} \end{bmatrix} \in \mathcal{L}_1^{\text{loc}}\left(\mathbb{R}, \mathbb{R}^{n+d}\right) \,\middle|\, \begin{array}{l} \frac{d\mathbf{x}}{dt} = A\mathbf{x} + B\mathbf{u}, \text{ and } \mathbf{y} = C\mathbf{x} + D(\frac{d}{dt})\mathbf{u} \\ A \in \mathbb{R}^{d \times d}, B \in \mathbb{R}^{d \times (n-m)}, \\ C \in \mathbb{R}^{m \times d}, \text{ and } D \in \mathbb{R}^{m \times (n-m)}[\xi] \end{array} \right\}. \tag{2.3}$$

To determine whether $\mathcal{B}_s$ realizes $\mathcal{B}_{i/o}$, we must eliminate $\mathbf{x}$ from the equations:

$$\bar{R}(\tfrac{d}{dt})\text{col}\begin{pmatrix} \mathbf{y} & \mathbf{u} & \mathbf{x} \end{pmatrix} = 0, \text{ where } \bar{R} := \begin{bmatrix} I_m & -D & -C \\ 0 & B & -\mathcal{A} \end{bmatrix}, \text{ with } \mathcal{A}(\xi) = \xi I_d - A. \tag{2.4}$$

Elimination of variables is enabled by the non-uniqueness of the representation of behaviors. From [17, Theorem 3.6.2], if $R, \tilde{R} \in \mathbb{R}^{l \times n}[\xi]$, then $\mathcal{B}$ in (2.1) satisfies $\mathcal{B} = \{\mathbf{w} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \mid \tilde{R}(\frac{d}{dt})\mathbf{w} = 0\}$ if and only if there exists a unimodular U such that $R = U\tilde{R}$. For $\bar{R}$ in (2.4), since $\det(\mathcal{A})$ is the characteristic polynomial of A, then the final d columns of $\bar{R}(\lambda)$ are independent (so $\bar{R}(\lambda)$ has full row rank) for almost all $\lambda \in \mathbb{C}$. Following [17, Theorem 6.2.6], we obtain a relationship of the form

$$\begin{bmatrix} U_{1,1} & U_{1,2} \\ U_{2,1} & U_{2,2} \end{bmatrix} \begin{bmatrix} I_m & -D & -C \\ 0 & B & -\mathcal{A} \end{bmatrix} = \begin{bmatrix} \tilde{R}_1 & -\tilde{R}_2 & 0 \\ Z_1 & Z_2 & -Z_3 \end{bmatrix}, \tag{2.5}$$

where the leftmost matrix is unimodular, and where $Z_3 \in \mathbb{R}^{d\times d}[\xi]$ and $Z_3(\lambda)$ is nonsingular for almost all $\lambda \in \mathbb{C}$. We note that this relationship can be obtained by computing an upper-echelon or row reduced form for $\operatorname{col}\begin{pmatrix} C & \mathcal{A}\end{pmatrix}$ (see Appendix A). As will be shown in Theorem 5.2 of this paper, $\mathbf{x}$ is *properly eliminable* in $\mathcal{B}_s$ (see [16]), which implies that $\{\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R},\mathbb{R}^n) \mid \exists \mathbf{x} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R},\mathbb{R}^d)$ satisfying (2.4)$\}$ is equal to $\{\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R},\mathbb{R}^n) \mid \tilde{R}_1(\frac{d}{dt})\mathbf{y} = \tilde{R}_2(\frac{d}{dt})\mathbf{u}\}$ [16]. We conclude that $\mathcal{B}_{i/o}$ in (2.2) satisfies $\mathcal{B}_{i/o} = \{\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \mid \exists \mathbf{x}$ with $\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u} & \mathbf{x}\end{pmatrix} \in \mathcal{B}_s\}$ if and only if there exists a unimodular W such that $W\begin{bmatrix}\tilde{R}_1 & -\tilde{R}_2\end{bmatrix} = \begin{bmatrix}R_1 & -R_2\end{bmatrix}$.

Finally, $\mathcal{B}$ in (2.1) is called behaviorally controllable if, for any $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{B}$, there exists a $t_1 \geq 0$ and a $\mathbf{w} \in \mathcal{B}$ which satisfies $\mathbf{w}(t) = \mathbf{w}_1(t)$ for all $t \leq 0$ and $\mathbf{w}(t) = \mathbf{w}_2(t)$ for all $t \geq t_1$ [17, Definition 5.2.2]. From [17, Theorem 5.2.10], $\mathcal{B}$ is behaviorally controllable if and only if the rank of $R(\lambda)$ is the same for all $\lambda \in \mathbb{C}$.

3. Realization of transfer functions and behaviors. Realization theory for linear systems is typically associated with the realization of transfer functions: given $G \in \mathbb{R}^{m\times p}(\xi)$, find $A \in \mathbb{R}^{\bullet\times\bullet}$, $B \in \mathbb{R}^{\bullet\times p}$, $C \in \mathbb{R}^{m\times\bullet}$, and $D \in \mathbb{R}^{m\times p}[\xi]$, such that

$$G(\xi) = D(\xi) + C(\xi I - A)^{-1}B. \tag{3.1}$$

Of particular significance are *minimal* realizations which have the additional properties that the pair (A, B) is *controllable* and the pair (C, A) is *observable*, where

$$(A, B) \text{ is controllable } \iff \begin{bmatrix}B & \lambda I - A\end{bmatrix} \text{ has full row rank for all } \lambda \in \mathbb{C}, \text{ and} \tag{3.2}$$

$$(C, A) \text{ is observable } \iff \operatorname{col}\begin{pmatrix}C & \lambda I - A\end{pmatrix} \text{ has full column rank for all } \lambda \in \mathbb{C}. \tag{3.3}$$

The first general solution to this problem appeared in [13]. This was followed by solutions based on the Markov parameters for G, which are the terms in the formal series expansion $C(\xi I - A)^{-1}B = CB/\xi + CAB/(\xi^2) + CA^2B/(\xi^3) + \ldots$ [9, 22]. If $G = R_1^{-1}R_2$ where $R_1 \in \mathbb{R}^{m\times m}[\xi]$ and $R_2 \in \mathbb{R}^{m\times(n-m)}[\xi]$ are coprime, and (3.1) is a minimal realization of G, then the behavior $\mathcal{B}_{i/o}$ in (2.2) satisfies $\mathcal{B}_{i/o} = \{\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \mid \exists \mathbf{x}$ with $\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u} & \mathbf{x}\end{pmatrix} \in \mathcal{B}_s\}$, for $\mathcal{B}_s$ as in (2.3). Whenever R_1 and R_2 are coprime, then $\begin{bmatrix}R_1(\lambda) & -R_2(\lambda)\end{bmatrix}$ has full row rank for all $\lambda \in \mathbb{C}$, and hence $\mathcal{B}_{i/o}$ is behaviorally controllable (see Section 2). However, these realization procedures are unable to represent uncontrollable behaviors, as the following example will demonstrate.

We consider the driving-point behavior of the network in Fig. 1:

$$\mathcal{B}_d := \left\{ \begin{bmatrix} v \\ i \end{bmatrix} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R},\mathbb{R}^2) \,\middle|\, (\tfrac{d}{dt}+1)\begin{bmatrix}(\tfrac{d^2}{dt^2}+\tfrac{d}{dt}+4) & -(\tfrac{d^2}{dt^2}+\tfrac{d}{dt}+1)\end{bmatrix}\begin{bmatrix} v \\ i \end{bmatrix} = 0 \right\}. \tag{3.4}$$

Following [22], to obtain a realization of the transfer function $(\xi^2+\xi+1)/(\xi^2+\xi+4)$, we consider the Markov parameters $H_0, H_1, H_2, \ldots$ in the formal series expansion $(\xi^2+\xi+1)/(\xi^2+\xi+4) = 1 + 0/\xi - 3/\xi^2 + \ldots =: H_0 + H_1/\xi + H_2/\xi^2 + \ldots$. By multiplying through by $\xi^2+\xi+4$ and then equating coefficients of ξ^{-k}, we find that $H_{k+2} = -H_{k+1} - 4H_k$ for $k \geq 1$. Thus, with

$$\hat{A} = \begin{bmatrix}-1 & 1 \\ -4 & 0\end{bmatrix},\ \hat{B} = \begin{bmatrix}0 \\ 1\end{bmatrix},\ \hat{C} = \begin{bmatrix}H_2 & H_1\end{bmatrix} = \begin{bmatrix}-3 & 0\end{bmatrix}, \text{ and } \hat{D} = H_0 = 1 \tag{3.5}$$

we have $\hat{C}\hat{A}^k = \begin{bmatrix}H_{k+2} & H_{k+1}\end{bmatrix}$, and hence $\hat{C}\hat{A}^k\hat{B} = H_{k+1}$ $(k = 0, 1, \ldots)$. It follows that $(\xi^2+\xi+1)/(\xi^2+\xi+4) = H_0 + H_1/\xi + H_2/(\xi^2) + \ldots = D + \hat{C}(\xi I - \hat{A})^{-1}\hat{B}$. Now, consider $\operatorname{col}\begin{pmatrix}\hat{v} & \hat{i}\end{pmatrix}$ with $\hat{v}(t) = e^{-t}$ and $\hat{i}(t) = 0$ for all $t \in \mathbb{R}$. Then $\operatorname{col}\begin{pmatrix}\hat{v} & \hat{i}\end{pmatrix} \in \mathcal{B}_d$. However, if $\mathbf{x} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R},\mathbb{R}^2)$ satisfies $\frac{d\mathbf{x}}{dt} = \hat{A}\mathbf{x} + \hat{B}\hat{i} = \hat{A}\mathbf{x}$, then $\mathbf{x}(t) =$

$\alpha\text{col}\left(\cos(\sqrt{15}t/2) \quad 2\cos(\sqrt{15}t/2+\phi)\right) + \beta\text{col}\left(\sin(\sqrt{15}t/2) \quad 2\sin(\sqrt{15}t/2+\phi)\right)$ for some $\alpha, \beta \in \mathbb{R}$, with $\phi := \arctan(\sqrt{15})$ (see [17, Theorem 4.5.17]), and we conclude that $\text{col}\left(\hat{v} \quad \hat{i}\right)$ cannot be written in the form $\hat{v} = \hat{C}\mathbf{x} + \hat{D}\hat{i}$ with $\frac{d\mathbf{x}}{dt} = \hat{A}\mathbf{x} + \hat{B}\hat{i}$.

In contrast, from Theorem 5.3 of this paper, we obtain a realization for $\mathcal{B}_d$ with

$$\hat{A} = \begin{bmatrix} -2 & 1 & 0 \\ -5 & 0 & 1 \\ -4 & 0 & 0 \end{bmatrix}, \; \hat{B} = \begin{bmatrix} 0 \\ -3 \\ -3 \end{bmatrix}, \; \hat{C} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \text{ and } \hat{D} = 1. \tag{3.6}$$

In this case, $\hat{v} = \hat{C}\hat{\mathbf{x}} + \hat{D}\hat{i}$ with $\frac{d\hat{\mathbf{x}}}{dt} = \hat{A}\hat{\mathbf{x}} + \hat{B}\hat{i}$ when $\hat{x}(t) = \text{col}\left(1 \quad 1 \quad 4\right)e^{-t}$ $(t \in \mathbb{R})$.

We note that the realizations in (3.5) and (3.6) correspond to the *controllability* and *observer* canonical forms for the single-input single-output system (3.4), respectively, and both incorporate companion matrices. As shown in [15, 14, 6, 1], the properties of companion matrices facilitate efficient computation. While many realization procedures incorporate companion matrices in the single-input single-output case, we are unaware of any procedures which also incorporate companion matrices in the multi-input multi-output case, as is the case for the realizations presented in Theorems 4.1 and 5.3 of this paper (we note that there are procedures incorporating block companion matrices, but these prohibit the application of the results in [15, 14, 6, 1]). The advantages of this are demonstrated in Section 6, where we exploit the properties of companion matrices to efficiently compute trajectories, and the solution to Lyapunov equations, for our realizations.

To conclude this section, we compare our approach to other solutions to the behavioral realization problem. Firstly, in [20, proof of Theorem 3], a realization is provided for the behavior of a discrete time system analogous to $\mathcal{B}_{i/o}$ in (2.2), providing $R := \begin{bmatrix} R_1 & -R_2 \end{bmatrix}$ is in *row reduced form* (see Section 7). Secondly, the papers [18, 19] consider a behavior $\mathcal{B}$ as in (2.1) for which R is in row reduced form, and the primary focus is the construction of a *state map* $X\left(\frac{d}{dt}\right)$, $X \in \mathbb{R}^{d \times n}[\xi]$, and real matrices E, F, G, such that $\mathcal{B} = \{\mathbf{w} \mid \exists \mathbf{x} \text{ with col}\left(\mathbf{w} \quad \mathbf{x}\right) \in \mathcal{B}_f\}$, where $\mathbf{x} = X\left(\frac{d}{dt}\right)\mathbf{w}$, and $\mathcal{B}_f = \{\text{col}\left(\mathbf{w} \quad \mathbf{x}\right) \in \mathcal{L}_1^{\text{loc}}\left(\mathbb{R}, \mathbb{R}^{n+d}\right) \mid E\frac{d\mathbf{x}}{dt} + F\mathbf{x} + G\mathbf{w} = 0\}$. The construction of analogous representations for discrete time systems was considered in [7].

As in [20], Theorems 4.1 and 5.3 of this paper provide a realization for the behavior $\mathcal{B}_{i/o}$ in (2.2) (note that we do not require $R := \begin{bmatrix} R_1 & -R_2 \end{bmatrix}$ to be in row reduced form), and our realizations can be efficiently computed using existing algorithms in symbolic algebra programs. Theorems 4.1 and 5.3 also yield realizations for $\mathcal{B}$ in (2.1) using the results in Section 7. Most importantly, unlike the realizations in [20, 18, 19, 7], our realizations incorporate companion matrices in the multi-input multi-output case. This is advantageous for analysis and simulation as discussed earlier in this section.

4. Behavioral realizations using companion matrices. The main result in this section is Theorem 4.1, which provides a realization for the behavior $\mathcal{B}_{i/o}$ in (2.2) which incorporates companion matrices. In Section 6, we will show how to efficiently compute trajectories, and solutions to Lyapunov equations, for this realization.

The terms in Theorem 4.1 relate to the Smith form for R_1 as follows. Given the non-singularity of $R_1(\lambda)$ for almost all $\lambda \in \mathbb{C}$, the Smith form for R_1 leads to unimodular matrices $U, V \in \mathbb{R}^{m \times m}[\xi]$ and a diagonal $S \in \mathbb{R}^{m \times m}[\xi]$ such that $UR_1V = S$ (see Appendix A). Here, $S = \text{diag}\left(\sigma_1 \quad \cdots \quad \sigma_m\right)$, and each term in the sequence $\sigma_1, \ldots \sigma_m$ is non-zero and is divisible by the preceding term. Then, with $F^T := UR_2$, it follows that there exists a $0 \leq q \leq m$ and a unimodular $\hat{V} \in \mathbb{R}^{m \times m}[\xi]$ such that

$$U\begin{bmatrix} R_1 & -R_2 \end{bmatrix} = \begin{bmatrix} S\hat{V}^T & -F^T \end{bmatrix}, \text{ with } S = \text{diag}\left(I_{m-q} \quad \mu_1 \quad \cdots \quad \mu_q\right), \; \hat{V}^T = V^{-1}, \tag{4.1}$$

and where the degree of μ_j is equal to $d_j \geq 1$ $(j = 1, \ldots, q)$. We then define $\mu_{j,0}, \mu_{j,1}, \ldots, \mu_{j,d_j-1} \in \mathbb{R}$ as the coefficients in μ_j:

$$\mu_j(\xi) =: \xi^{d_j} + \mu_{j,d_j-1}\xi^{d_j-1} + \ldots + \mu_{j,1}\xi + \mu_{j,0}, \;\; j = 1, \ldots, q. \tag{4.2}$$

Next, we partition $V, \hat{V}$, and F compatibly with S as follows:

$$V =: \begin{bmatrix} V_1 & \mathbf{v}_1 & \cdots & \mathbf{v}_q \end{bmatrix}, \hat{V} =: \begin{bmatrix} \hat{V}_1 & \hat{\mathbf{v}}_1 & \cdots & \hat{\mathbf{v}}_q \end{bmatrix}, \text{and } F =: \begin{bmatrix} F_1 & \mathbf{f}_1 & \cdots & \mathbf{f}_q \end{bmatrix},$$

where $V_1, \hat{V}_1 \in \mathbb{R}^{m \times (m-q)}[\xi]$, $F_1 \in \mathbb{R}^{(n-m)\times(m-q)}[\xi]$, $\mathbf{v}_j, \hat{\mathbf{v}}_j \in \mathbb{R}^m[\xi]$, and $\mathbf{f}_j \in \mathbb{R}^{n-m}[\xi]$ $(j = 1, \ldots, q)$. In particular, since $\hat{V}^T = V^{-1}$, then

$$V_1\hat{V}_1^T + \sum_{j=1}^{q} \mathbf{v}_j\hat{\mathbf{v}}_j^T = I_m. \tag{4.3}$$

We now define $\mathbf{a}_j \in \mathbb{R}^{n-m}[\xi]$ and $\mathbf{b}_j \in \mathbb{R}^{n-m}[\xi]$ (resp. $\mathbf{g}_j \in \mathbb{R}^m[\xi]$ and $\mathbf{c}_j \in \mathbb{R}^m[\xi]$) as the quotient and remainder of $\mathbf{f}_j$ (resp. $\mathbf{v}_j$) on division by μ_j, respectively:[1]

$$\mathbf{f}_j =: \mathbf{a}_j\mu_j + \mathbf{b}_j, \;\; \text{and } \mathbf{v}_j =: \mathbf{g}_j\mu_j + \mathbf{c}_j, \;\; (j = 1, \ldots, q), \tag{4.4}$$

and so the degrees of both $\mathbf{b}_j$ and $\mathbf{c}_j$ are less than d_j. Accordingly, we define $\hat{\mathbf{b}}_{j,k} \in \mathbb{R}^{n-m}$ (resp. $\mathbf{c}_{j,k} \in \mathbb{R}^m$) for $k = 0, 1, \ldots, d_j - 1$ as the coefficients in $\mathbf{b}_j$ (resp. $\mathbf{c}_j$):

$$\begin{aligned} \mathbf{b}_j(\xi) &=: \hat{\mathbf{b}}_{j,0} + \hat{\mathbf{b}}_{j,1}\xi + \hat{\mathbf{b}}_{j,2}\xi^2 + \ldots + \hat{\mathbf{b}}_{j,d_j-1}\xi^{d_j-1}, \\ \text{and } \;\; \mathbf{c}_j(\xi) &=: \mathbf{c}_{j,0} + \mathbf{c}_{j,1}\xi + \mathbf{c}_{j,2}\xi^2 + \ldots + \mathbf{c}_{j,d_j-1}\xi^{d_j-1}, \;\; (j = 1, \ldots, q), \end{aligned}$$

and we define $\hat{B}_j \in \mathbb{R}^{d_j \times (n-m)}$ and $C_j \in \mathbb{R}^{m \times d_j}$ for $j = 1, \ldots, q$ as:

$$\hat{B}_j := \begin{bmatrix} \hat{\mathbf{b}}_{j,d_j-1} & \hat{\mathbf{b}}_{j,d_j-2} & \cdots & \hat{\mathbf{b}}_{j,0} \end{bmatrix}^T, \;\; \text{and } C_j := \begin{bmatrix} \mathbf{c}_{j,0} & \mathbf{c}_{j,1} & \cdots & \mathbf{c}_{j,d_j-1} \end{bmatrix}. \tag{4.5}$$

We then let $\mathbf{b}_{j,1}, \mathbf{b}_{j,2}, \ldots \in \mathbb{R}^{n-m}$ be the Markov parameters for $\mathbf{b}_j/\mu_j$. These are the terms in the formal series expansion:

$$\mathbf{b}_j(\xi)/\mu_j(\xi) =: \mathbf{b}_{j,1}\xi^{-1} + \mathbf{b}_{j,2}\xi^{-2} + \mathbf{b}_{j,3}\xi^{-3} + \ldots, \;\; (j = 1, \ldots, q). \tag{4.6}$$

By multiplying both sides of the above equation by μ_j and then equating coefficients, we obtain the matrix relationship:

$$Q_jB_j = \hat{B}_j, \tag{4.7}$$

$$\text{where } \;\; B_j := \begin{bmatrix} \mathbf{b}_{j,1} & \mathbf{b}_{j,2} & \mathbf{b}_{j,3} & \cdots & \mathbf{b}_{j,d_j} \end{bmatrix}^T, \tag{4.8}$$

$$\text{and } \;\; Q_j := \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ \mu_{j,d_j-1} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_{j,2} & \mu_{j,3} & \cdots & 1 & 0 \\ \mu_{j,1} & \mu_{j,2} & \cdots & \mu_{j,d_j-1} & 1 \end{bmatrix}, \;\; (j = 1, \ldots, q). \tag{4.9}$$

We further let A_j be the companion matrix and $\mathcal{A}_j$ the polynomial matrix:

$$A_j := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\mu_{j,0} & -\mu_{j,1} & -\mu_{j,2} & \cdots & -\mu_{j,d_j-1} \end{bmatrix}, \;\; \text{and } \mathcal{A}_j(\xi) := \xi I_{d_j} - A. \tag{4.10}$$

[1] Note that it is inefficient to directly compute the quotient and remainder for each entry in these vectors. Instead, it is better to compute the quotient and remainder for the monomial s^k from the quotient and remainder for s^{k-1} $(k = 1, 2, \ldots)$, and then take the appropriate linear sum.

Finally, we define $D \in \mathbb{R}^{m\times(n-m)}[\xi]$ as

$$D := V_1 F_1^T + \sum_{j=1}^{q} \mathbf{g}_j \mathbf{f}_j^T + \mathbf{c}_j \mathbf{a}_j^T + C_j P_j B_j, \tag{4.11}$$

$$\text{where} \quad P_j(\xi) := \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \xi & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \xi^{d_j-2} & \xi^{d_j-3} & \cdots & 1 & 0 \end{bmatrix}, \ (j = 1, \ldots, q). \tag{4.12}$$

THEOREM 4.1. *The behavior $\mathcal{B}_{i/o}$ in (2.2) has the realization:*

$$\mathcal{B}_{i/o} = \{col\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \mid \exists \mathbf{x} \text{ with } col\begin{pmatrix}\mathbf{y} & \mathbf{u} & \mathbf{x}\end{pmatrix} \in \mathcal{B}_s\}, \tag{4.13}$$

for $\mathcal{B}_s$ as in (2.3), with

$$A = diag\begin{pmatrix}A_1 & \cdots & A_q\end{pmatrix}, \ B = col\begin{pmatrix}B_1 & \cdots & B_q\end{pmatrix}, \ C = \begin{bmatrix}C_1 & \cdots & C_q\end{bmatrix},$$

and where $D, A_j, B_j,$ and C_j $(j = 1, \ldots, q)$ are defined in equations (4.1) to (4.12).

Prior to proving Theorem 4.1, we consider $\mathcal{B}_{i/o}$ in (2.2), and we let the Smith form for R_1 be $UR_1V = S$, and $F^T := UR_2$, where $S = \text{diag}\begin{pmatrix}1 & (s+1)^2 & (s+1)^2(s+2)\end{pmatrix}$,

$$U = \begin{bmatrix} 1 & 0 & 0 \\ -(s+1)^3 & 1 & 0 \\ (s+1)^2(s+2)(1+s-s^2) & -s-2 & 1 \end{bmatrix}, \ V = \begin{bmatrix} 1 & s(s^2-1) & s(s^2+s-1) \\ 0 & -s & -s-1 \\ 0 & 1 & 1 \end{bmatrix}, \ F^T = \begin{bmatrix} 2s(s-1) & 1 \\ s(s^2+3s+1) & -1 \\ 1-2s-5s^2-4s^3-s^4 & s \end{bmatrix}.$$

which was obtained using the Maple command `SmithForm`. Thus, $q = 2$, $\mu_1(\xi) = (\xi+1)^2$, $\mu_2(\xi) = (\xi+1)^2(\xi+2)$, $d_1 = 2$, and $d_2 = 3$. It follows that $V_1, \mathbf{v}_1$, and $\mathbf{v}_2$ are the first, second, and third columns of V, respectively. Also, $F_1, \mathbf{f}_1$, and $\mathbf{f}_2$ are the first, second, and third rows of F^T, respectively. Then, using the Maple command `RightDivision`, we obtain

$$\mathbf{g}_1 = \begin{bmatrix} s-2 \\ 0 \\ 0 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 2s+2 \\ -s \\ 1 \end{bmatrix}, \mathbf{g}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -3s^2-6s-2 \\ -1-s \\ 1 \end{bmatrix}, \mathbf{a}_1 = \begin{bmatrix} s+1 \\ 0 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} -2s-1 \\ -1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} -s \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ s \end{bmatrix}.$$

Then $\hat{B}_1$ and C_1 are readily obtained from the coefficients of $\mathbf{b}_1$ and $\mathbf{c}_1$, respectively; B_1 may subsequently be obtained from (4.7) by using the Maple command `ForwardSubstitute`; and D follows from (4.11). We thus obtain

$$A_1 = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}, B_1 = \begin{bmatrix} -2 & 0 \\ 3 & -1 \end{bmatrix}, C_1 = \begin{bmatrix} 2 & 2 \\ 0 & -1 \\ 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & -4 \end{bmatrix}, C_2 = \begin{bmatrix} -2 & -6 & -3 \\ -1 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, D = \begin{bmatrix} -1 & 0 \\ 2 & 0 \\ 1 & 0 \end{bmatrix}.$$

Proof (Theorem 4.1). To prove the present theorem, we must show that $\mathbf{x}$ is properly eliminable in $\mathcal{B}_s$ (this will be shown in Theorem 5.2), and we must demonstrate a relationship of the form of (2.5) in which the leftmost matrix is unimodular; $Z_3(\lambda)$ is non-singular for almost all $\lambda \in \mathbb{C}$; $\tilde{R}_1 = R_1$; and $\tilde{R}_2 = R_2$ (see Section 2). To obtain such a relationship, we first define $\mathbf{p}_j, \mathbf{q}_j \in \mathbb{R}^{d_j}[\xi]$, and $\mathbf{e}_j \in \mathbb{R}^{d_j}$, as

$$\mathbf{p}_j(\xi) := \text{col}\begin{pmatrix}1 & \xi & \cdots & \xi^{d_j-2} & \xi^{d_j-1}\end{pmatrix}, \tag{4.14}$$

$$\mathbf{q}_j(\xi) := \text{col}\begin{pmatrix}\xi^{d_j-1} & \xi^{d_j-2} & \cdots & \xi & 1\end{pmatrix}, \tag{4.15}$$

$$\text{and} \ \ \mathbf{e}_j := \text{col}\begin{pmatrix}0 & 0 & \cdots & 0 & 1\end{pmatrix}, \ (j = 1, \ldots, q). \tag{4.16}$$

We then let

$$X_{j,1} := \mathbf{p}_j \hat{\mathbf{v}}_j^T, \ X_{j,2} := -\mathbf{p}_j \mathbf{a}_j^T - P_j B_j, \ \ (j = 1, \ldots, q), \tag{4.17}$$

$$X_1 := \text{col}\begin{pmatrix}X_{1,1} & \cdots & X_{q,1}\end{pmatrix}, \ \text{and} \ X_2 := \text{col}\begin{pmatrix}X_{1,2} & \cdots & X_{q,2}\end{pmatrix}. \tag{4.18}$$

Then, with $d := \sum_{j=1}^{q} d_j$, we will show that

$$\begin{bmatrix} W_1 & C \\ W_2 & \mathcal{A} \end{bmatrix} \begin{bmatrix} S\hat{V}^T & -F^T & 0 \\ X_1 & X_2 & -I_d \end{bmatrix} = \begin{bmatrix} I_m & -D & -C \\ 0 & B & -\mathcal{A} \end{bmatrix}, \tag{4.19}$$

$$\text{where } W_1 := \begin{bmatrix} V_1 & \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_q \end{bmatrix}, \tag{4.20}$$

$$\text{and } W_2 := -\begin{bmatrix} 0_{d\times(m-q)} & \operatorname{diag}\begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_q \end{pmatrix} \end{bmatrix}. \tag{4.21}$$

It then follows from (4.1) and (4.19) that

$$Y \begin{bmatrix} R_1 & -R_2 & 0 \\ X_1 & X_2 & -I_d \end{bmatrix} = \begin{bmatrix} I_m & -D & -C \\ 0 & B & -\mathcal{A} \end{bmatrix}, \text{ with } Y := \begin{bmatrix} W_1 & C \\ W_2 & \mathcal{A} \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & I_d \end{bmatrix}. \tag{4.22}$$

Finally, we will show that $\det(Y) = \det(U)\det(V)$, and we conclude that Y is unimodular since U and V are. This will complete the proof.

To demonstrate the equality (4.19), we note initially that the final d block columns on the left and right hand sides of (4.19) are clearly equivalent. Therefore, it suffices to show the following four relationships: (i) $W_1S\hat{V}^T + CX_1 = I_m$, (ii) $W_2S\hat{V}^T + \mathcal{A}X_1 = 0$, (iii) $-W_1F^T + CX_2 = -D$, and (iv) $-W_2F^T + \mathcal{A}X_2 = B$. First, note that

$$\mathcal{A}_j\mathbf{p}_j = \mathbf{e}_j\mu_j, \text{ and } \mathcal{A}_jP_j = \mathbf{e}_j\mathbf{q}_j^TQ_j - I_{d_j}, \quad (j = 1, \ldots, q), \tag{4.23}$$

where the latter relationship may be verified by considering each row of $\mathcal{A}_jP_j$ in turn. Furthermore, it is clear that

$$\mathbf{q}_j^T\hat{B}_j = \mathbf{b}_j^T, \text{ and } C_j\mathbf{p}_j = \mathbf{c}_j, \quad (j = 1, \ldots, q). \tag{4.24}$$

Then, to see (i), note that the partitions of the matrices S and $\hat{V}$ imply $W_1S\hat{V}^T = V_1\hat{V}_1^T + \sum_{j=1}^{q} \mathbf{g}_j\mu_j\hat{\mathbf{v}}_j^T$, and that $CX_1 = \sum_{j=1}^{q} C_j\mathbf{p}_j\hat{\mathbf{v}}_j$ by (4.17). From (4.3), (4.4), and (4.24), we obtain $W_1S\hat{V}^T + CX_1 = V_1\hat{V}_1^T + \sum_{j=1}^{q}(\mathbf{g}_j\mu_j + \mathbf{c}_j)\hat{\mathbf{v}}_j^T = I_m$.

For (ii), note that $W_2S\hat{V}^T = -\operatorname{diag}\begin{pmatrix}\mathbf{e}_1 & \cdots & \mathbf{e}_q\end{pmatrix}\operatorname{col}\begin{pmatrix}\mu_1\hat{\mathbf{v}}_1^T & \cdots & \mu_q\hat{\mathbf{v}}_q^T\end{pmatrix} = -\operatorname{col}\begin{pmatrix}\mathbf{e}_1\mu_1\hat{\mathbf{v}}_1^T & \cdots & \mathbf{e}_q\mu_q\hat{\mathbf{v}}_q^T\end{pmatrix}$, and $\mathcal{A}X_1 = \operatorname{diag}\begin{pmatrix}\mathcal{A}_1 & \cdots & \mathcal{A}_q\end{pmatrix}\operatorname{col}\begin{pmatrix}X_{1,1} & \cdots & X_{q,1}\end{pmatrix} = \operatorname{col}\begin{pmatrix}\mathcal{A}_1\mathbf{p}_1\hat{\mathbf{v}}_1^T & \cdots & \mathcal{A}_q\mathbf{p}_q\hat{\mathbf{v}}_q^T\end{pmatrix}$ by (4.17). Thus, $W_2S\hat{V}^T + \mathcal{A}X_1 = 0$ by (4.23).

To see (iii), observe that $-W_1F^T = -V_1F_1^T - \sum_{j=1}^{q}\mathbf{g}_j\mathbf{f}_j^T$, and that $CX_2 = -\sum_{j=1}^{q}\left(C_j\mathbf{p}_j\mathbf{a}_j^T + C_jP_jB_j\right)$ by (4.17). We then find that $-W_1F^T + CX_2 = -D$ from (4.11) and (4.24).

Finally, for (iv), note that $-W_2F^T = \operatorname{diag}\begin{pmatrix}\mathbf{e}_1 & \cdots & \mathbf{e}_q\end{pmatrix}\operatorname{col}\begin{pmatrix}\mathbf{f}_1^T & \cdots & \mathbf{f}_q^T\end{pmatrix} = \operatorname{col}\begin{pmatrix}\mathbf{e}_1\mathbf{f}_1^T & \cdots & \mathbf{e}_q\mathbf{f}_q^T\end{pmatrix}$, and that $\mathcal{A}X_2 = \operatorname{diag}\begin{pmatrix}\mathcal{A}_1 & \cdots & \mathcal{A}_q\end{pmatrix}\operatorname{col}\begin{pmatrix}X_{1,2} \cdots & X_{q,2}\end{pmatrix} = -\operatorname{col}\begin{pmatrix}\mathcal{A}_1(\mathbf{p}_1\mathbf{a}_1^T + P_1B_1) & \cdots & \mathcal{A}_q(\mathbf{p}_q\mathbf{a}_q^T + P_qB_q)\end{pmatrix}$ by (4.17). Furthermore, from (4.4), (4.7), (4.23) and (4.24), we find that $\mathbf{e}_j\mathbf{f}_j^T - \mathcal{A}_j(\mathbf{p}_j\mathbf{a}_j^T + P_jB_j) = \mathbf{e}_j\mu_j\mathbf{a}_j^T + \mathbf{e}_j\mathbf{b}_j^T - \mathbf{e}_j\mu_j\mathbf{a}_j^T - \mathbf{e}_j\mathbf{q}_j^TQ_jB_j + B_j = \mathbf{e}_j\mathbf{q}_j^T(\hat{B}_j - Q_jB_j) + B_j = B_j$. Thus, $-W_2F^T + \mathcal{A}X_2 = \operatorname{col}\begin{pmatrix}B_1 & \cdots & B_q\end{pmatrix} = B$.

It remains to show that $\det(Y) = \det(U)\det(V)$. Firstly, note that $\det(\mathcal{A}) = \prod_{j=1}^{q}\det(\mathcal{A}_j) = \prod_{j=1}^{q}\mu_j = \det(S)$ [8, p. 149]. In particular, $\mathcal{A}$ is non-singular, and

$$\begin{bmatrix} I & -C\mathcal{A}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} W_1 & C \\ W_2 & \mathcal{A} \end{bmatrix} = \begin{bmatrix} W_1 - C\mathcal{A}^{-1}W_2 & 0 \\ W_2 & \mathcal{A} \end{bmatrix}. \tag{4.25}$$

By pre-multiplying both sides of (4.19) by the leftmost matrix in (4.25) and comparing the top left blocks in the resulting equation, we find that $(W_1 - C\mathcal{A}^{-1}W_2)S\hat{V}^T = I_m$, whence $(W_1 - C\mathcal{A}^{-1}W_2)S = V$ by (4.1). Since the determinant of the leftmost matrix in (4.25) is equal to one, then combining the preceding relationships with (4.22) gives $\det(Y) = \det(U)\det(W_1 - C\mathcal{A}^{-1}W_2)\det(\mathcal{A}) = \det(U)\det(W_1 - C\mathcal{A}^{-1}W_2)\det(S) = \det(U)\det(V)$. ∎

5. Controllability, observability, dimension, and proper elimination. In this section, we demonstrate several properties of the realization in Theorem 4.1. In particular, we show that it is observable; that it is controllable if and only if $\mathcal{B}_{i/o}$ is behaviorally controllable; and that it has the least possible *dimension*, equal to $\Delta(\mathcal{B}_{i/o})$ (see Theorem 5.2). We then provide a second realization with these same properties (Theorem 5.3). Here, $\Delta(\mathcal{B}_{i/o})$ is defined as follows:

DEFINITION 5.1. *Let $\mathcal{B} = \{\mathbf{w} \in \mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^n) \mid \hat{R}(\frac{d}{dt})\mathbf{w} = 0\}$ where $\hat{R} \in \mathbb{R}^{m \times n}$ and $\hat{R}(\lambda)$ has full row rank for almost all $\lambda \in \mathbb{C}$. Then $\Delta(\mathcal{B}) := \Delta(\hat{R})$, where $\Delta(\hat{R})$ denotes the maximum degree among all $m \times m$ minors of $\hat{R}$.*

Note that this definition uniquely assigns an integer $\Delta(\mathcal{B})$ to any behavior $\mathcal{B}$ of the form (2.1). This follows since $\mathcal{B}$ necessarily has a representation as in (7.2) (see Section 7). If, in addition, $\mathcal{B}$ also has the representation in Definition 5.1, then there exists a unimodular U such that $\hat{R} = U\tilde{R}$ [17, Theorem 3.6.4], whence $\Delta(\hat{R}) = \Delta(\tilde{R})$.

Prior to stating Theorem 5.2, we introduce some further notation. For $R \in \mathbb{R}^{m \times n}[\xi]$, we denote the minor formed from rows $q_1, \ldots, q_r$ and columns $p_1, \ldots, p_r$ of R by $R\left(\begin{smallmatrix} q_1, & \ldots, & q_r \\ p_1, & \ldots, & p_r \end{smallmatrix}\right)$, and we denote the maximum degree among all minors of R (of any size) by $\delta(R)$. We note that this is equal to the McMillan degree of R, since all poles of R are at infinity [22]. Further, providing the row rank of $R(\lambda)$ is m (i.e. $R(\lambda)$ has full row rank) for almost all $\lambda \in \mathbb{C}$, then we denote the minor formed from columns $p_1, \ldots, p_m$ of R ($p_1 < \ldots < p_m$) by $R(p_1, \ldots, p_m)$.

We now state Theorem 5.2. Note that the first part of this theorem (showing that $\mathbf{x}$ is properly eliminable in $\mathcal{B}_s$) is required in the proof of Theorem 4.1.

THEOREM 5.2. *Let $\mathcal{B}_{i/o}$ have the representation (4.13) with $\mathcal{B}_s$ as in (2.3). Then $\mathbf{x}$ is properly eliminable in $\mathcal{B}_s$, so $\mathcal{B}_{i/o}$ also takes the form (2.2) for some $R_1 \in \mathbb{R}^{m \times m}[\xi]$ and $R_2 \in \mathbb{R}^{(n-m) \times m}[\xi]$. Furthermore, D in (2.3) satisfies $R_1^{-1}R_2 = G + D$ with G strictly proper, and $\delta(D) + d \geq \Delta(\mathcal{B}_{i/o})$ (see Definition 5.1). If, in addition, $A, B, C,$ and D are as defined in Theorem 4.1, then the following conditions all hold:*

1. *(C, A) is observable.*
2. *(A, B) is controllable if and only if $\mathcal{B}_{i/o}$ is behaviorally controllable.*
3. *$\delta(D) + d = \Delta(\mathcal{B}_{i/o})$.*

Proof. Let $\bar{R}$ be as in (2.4). We will show that the maximum degree among all $(m+d) \times (m+d)$ minors of $\bar{R}$ (i.e., $\Delta(\bar{R})$) is $\delta(D) + d$, and that this degree is attained by an $(m+d) \times (m+d)$ minor of $\bar{R}$ which includes the d columns in $\operatorname{col}\begin{pmatrix} -C & -\mathcal{A} \end{pmatrix}$. That $\mathbf{x}$ is properly eliminable in $\mathcal{B}_s$ then follows from [16, Theorem 2.8].

We first define $\bar{R}_1 \in \mathbb{R}^{m \times (n+d)}[\xi]$ and $\bar{R}_2 \in \mathbb{R}^{d \times (n+d)}[\xi]$ as the matrices formed from the first m and last d rows of $\bar{R}$, respectively. In other words:

$$\bar{R}_1 := \begin{bmatrix} I_m & -D & -C \end{bmatrix}, \text{ and } \bar{R}_2 := \begin{bmatrix} 0 & B & -\mathcal{A} \end{bmatrix}. \tag{5.1}$$

Then, by expressing the $(m+d) \times (m+d)$ minor $\bar{R}(p_1, \ldots, p_{m+d})$ ($1 \leq p_1 < \ldots < p_{m+d} \leq n + d$) in terms of minors comprised of entries from each of the first m rows of R and their complementary minors, we obtain

$$\bar{R}(p_1, \ldots, p_{m+d}) = \sum_{\substack{l_1 < \ldots < l_m \in \{p_1, \ldots, p_{m+d}\} \\ k_1 < \ldots < k_d \in \{p_1, \ldots, p_{m+d}\} \setminus \{l_1, \ldots, l_m\}}} \bar{R}_1(l_1, \ldots, l_m) \times \bar{R}_2(k_1, \ldots, k_d) \times \epsilon(l_1, \ldots, l_m, k_1, \ldots, k_d), \tag{5.2}$$

with $\epsilon(l_1,\ldots,l_m,k_1,\ldots,k_d) = (-1)^n$ where n is the number of transpositions required to bring the sequence $l_1,\ldots,l_m,k_1,\ldots,k_d$ into numerical order. From (5.1), it is evident that $\deg\left(\bar{R}_2(k_1,\ldots,k_d)\right) \leq d$, with equality if and only if $k_j = n+j$ for $j = 1,\ldots,d$. Furthermore, by considering the expressions for the determinants $\bar{R}_1(l_1,\ldots,l_m)$ in terms of minors comprised exclusively of entries from those columns contained in $-D$ and complementary minors comprised exclusively of entries from those columns not contained in $-D$, it is evident that $\deg\left(\bar{R}_1(l_1,\ldots,l_m)\right) \leq \delta(D)$. Thus, from (5.2), we see that $\deg\left(\bar{R}(p_1,\ldots,p_{m+d})\right) \leq \delta(D)+d$.

Now, suppose the degree $\delta(D)$ is attained by a minor of D comprised of the columns $j_1,\ldots,j_\alpha$ and the rows $l_1,\ldots,l_\alpha$ of D ($1 \leq j_1 < \ldots < j_\alpha \leq n-m$, and $1 \leq l_1 < \ldots < l_\alpha \leq m$), and denote the remaining rows of D by $l_{\alpha+1},\ldots,l_m$ ($1 \leq l_{\alpha+1} < \ldots < l_m \leq m$). Then, in the expression for the determinant $\bar{R}_1(l_{\alpha+1},\ldots,l_m,m+j_1,\ldots,m+j_\alpha)$ in terms of the minors comprised of entries from each of the columns $j_1,\ldots,j_\alpha$ of $-D$ and their complementary minors, the only non-zero term is equal to

$$\pm\bar{R}_1\begin{pmatrix} l_{\alpha+1}, & \ldots, & l_m \\ l_{\alpha+1}, & \ldots, & l_m\end{pmatrix} \times \bar{R}_1\begin{pmatrix} l_1, & \ldots, & l_\alpha \\ m+j_1, & \ldots, & m+j_\alpha\end{pmatrix} = \pm D\begin{pmatrix} l_1, & \ldots, & l_\alpha \\ j_1, & \ldots, & j_\alpha\end{pmatrix}.$$

It thus follows that $\deg\left(\bar{R}_1(l_{\alpha+1},\ldots,l_m,m+j_1,\ldots,m+j_\alpha)\right) = \delta(D)$. Then, in the expansion (5.2) for $\bar{R}(l_{\alpha+1},\ldots,l_m,m+j_1,\ldots,m+j_\alpha,n+1,\ldots,n+d)$, all of the terms in the summation have degree strictly less that $\delta(D)+d$, with the exception of one term which is equal to $\pm\bar{R}_1(l_{\alpha+1},\ldots,l_m,m+j_1,\ldots,m+j_\alpha) \times \bar{R}_2(n+1,\ldots,n+d)$, and so has degree equal to $\delta(D)+d$.

Now, let $\mathcal{B}_{i/o}$ in (4.13) also satisfy (2.2). Then, from Section 2, there exists a relationship of the form of (2.5) in which (i) the leftmost matrix is unimodular; (ii) $Z_3(\lambda)$ is non-singular for almost all $\lambda \in \mathbb{C}$; and (iii) $W\begin{bmatrix}\tilde{R}_1 & -\tilde{R}_2\end{bmatrix} = \begin{bmatrix}R_1 & -R_2\end{bmatrix}$ for some unimodular W. To prove the inequality $\delta(D)+d \geq \Delta(\mathcal{B}_{i/o})$, we note that since $R_1(\lambda)$ in (4.13) is non-singular for almost all $\lambda \in \mathbb{C}$, then both $\begin{bmatrix}R_1(\lambda) & -R_2(\lambda)\end{bmatrix}$ and $\begin{bmatrix}\tilde{R}_1(\lambda) & -\tilde{R}_2(\lambda)\end{bmatrix}$ have full row rank for almost all $\lambda \in \mathbb{C}$. Moreover, from (iii), we have $\mathcal{B}_{i/o} = \{\mathrm{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \in \mathcal{L}_1^{\mathrm{loc}}(\mathbb{R},\mathbb{R}^n) \mid \begin{bmatrix}\tilde{R}_1 & -\tilde{R}_2\end{bmatrix}\left(\frac{d}{dt}\right)\mathrm{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} = 0\}$ (see Section 2). Then, from the preceding argument, we obtain

$$\begin{aligned}
(5.3)\qquad \delta(D)+d &= \max_{j_1<\ldots<j_m\in\{1,\ldots,n\}} \deg\left(\bar{R}(j_1,\ldots,j_m,n+1,\ldots,n+d)\right)\\
&= \max_{j_1<\ldots<j_m\in\{1,\ldots,n\}} \deg\left(\begin{bmatrix}\tilde{R}_1 & -\tilde{R}_2 & 0\\ Z_1 & Z_2 & -Z_3\end{bmatrix}(j_1,\ldots,j_m,n+1,\ldots,n+d)\right)\\
&= \Delta\left(\begin{bmatrix}\tilde{R}_1 & -\tilde{R}_2\end{bmatrix}\right) + \deg\left(\det\left(Z_3\right)\right) \geq \Delta(\mathcal{B}_{i/o}),
\end{aligned}$$

We next show that $R_1^{-1}R_2 = D+G$ with $G = C\mathcal{A}^{-1}B$, which is strictly proper. To see this, we recall that $W\begin{bmatrix}\tilde{R}_1 & -\tilde{R}_2\end{bmatrix} = \begin{bmatrix}R_1 & -R_2\end{bmatrix}$ with W unimodular and with $\tilde{R}_1$ and $\tilde{R}_2$ as in (2.5). As $R_1(\lambda)$ is non-singular for almost all $\lambda \in \mathbb{C}$, then so too is $\tilde{R}_1(\lambda)$, and hence $R_1^{-1}R_2 = \tilde{R}_1^{-1}\tilde{R}_2$. Since, in addition, $Z_3(\lambda)$ is non-singular for almost all $\lambda \in \mathbb{C}$, then (2.5) implies

$$\begin{bmatrix}\tilde{R}_1 & 0\\ Z_1 & -Z_3\end{bmatrix}^{-1}\begin{bmatrix}-\tilde{R}_2\\ Z_2\end{bmatrix} = \begin{bmatrix}I_m & -C\\ 0 & -\mathcal{A}\end{bmatrix}^{-1}\begin{bmatrix}-D\\ B\end{bmatrix} = \begin{bmatrix}I_m & -C\mathcal{A}^{-1}\\ 0 & -\mathcal{A}^{-1}\end{bmatrix}\begin{bmatrix}-D\\ B\end{bmatrix}.$$

Thus, from the first block row in the above equation, we obtain $R_1^{-1}R_2 = D+C\mathcal{A}^{-1}B$.

It remains to show conditions 1 to 3 when A, B, C, and D are as defined in Theorem 4.1. For condition 1, we recall relationships (4.19) to (4.21). In the proof of

Theorem 4.1 it was shown that the leftmost matrix in (4.19) is unimodular. Hence, $\text{col}\begin{pmatrix} C & \mathcal{A}(\lambda) \end{pmatrix}$ has full column rank for all $\lambda \in \mathbb{C}$, so (C, A) is observable by (3.3).

To see condition 2, we recall the relationship (4.22), and we denote

$$\hat{R} := \begin{bmatrix} R_1 & -R_2 & 0 \\ X_1 & X_2 & -I_d \end{bmatrix}, \text{ and } \bar{R} := \begin{bmatrix} I_m & -D & -C \\ 0 & B & -\mathcal{A} \end{bmatrix}. \tag{5.4}$$

It is then clear that $\bar{R}(\lambda)$ has full row rank for all $\lambda \in \mathbb{C}$ if and only if $\begin{bmatrix} B & -\mathcal{A}(\lambda) \end{bmatrix}$ does, and evidently $\begin{bmatrix} B & -\mathcal{A}(\lambda) \end{bmatrix}$ has full row rank for all $\lambda \in \mathbb{C}$ if and only if $\begin{bmatrix} B & \mathcal{A}(\lambda) \end{bmatrix}$ does. Similarly, $\hat{R}(\lambda)$ has full row rank for all $\lambda \in \mathbb{C}$ if and only if $\begin{bmatrix} R_1(\lambda) & -R_2(\lambda) \end{bmatrix}$ does. Furthermore, from the proof of Theorem 4.1, $Y\hat{R} = \bar{R}$ for Y as in (4.22), which is unimodular, and hence $\bar{R}(\lambda)$ has full row rank for all $\lambda \in \mathbb{C}$ if and only if $\hat{R}(\lambda)$ does. From (3.2), we conclude that (A, B) is controllable if and only if $\mathcal{B}_{i/o}$ is behaviorally controllable (see Section 2).

Finally, since $Y\hat{R} = \bar{R}$ with Y unimodular, then $Y^{-1}\bar{R} = \hat{R}$ with Y^{-1} unimodular, and we note that this takes the form of (2.5) by identifying Y^{-1} with the leftmost matrix in (2.5), and by identifying R_1, R_2, X_1, X_2, and I_d with $\tilde{R}_1, \tilde{R}_2, Z_1, Z_2$ and Z_3, respectively. Following the argument in the paragraph preceding equation (5.3), we conclude that $\delta(D) + d = \Delta\left(\begin{bmatrix} R_1 & -R_2 \end{bmatrix}\right) + \deg\left(\det\left(I_d\right)\right) = \Delta(\mathcal{B}_{i/o})$. ◻

We now present an alternative realization for $\mathcal{B}_{i/o}$. We recall that $\mathbf{a}_j$ and $\mathbf{b}_j$ (resp. $\mathbf{g}_j$ and $\mathbf{c}_j$) in (4.4) are the quotient and remainder of $\mathbf{f}_j$ (resp. $\mathbf{v}_j$) on division by μ_j, and we recall the relationship (4.7), where $B_j, \hat{B}_j$, and Q_j are as defined in Section 4. We now let $\hat{\mathbf{c}}_{j,1}, \hat{\mathbf{c}}_{j,2}, \ldots \in \mathbb{R}^m$ be the Markov parameters for $\mathbf{c}_j/\mu_j$. These are the terms in the formal series expansion:

$$\mathbf{c}_j(\xi)/\mu_j(\xi) = \hat{\mathbf{c}}_{j,1}\xi^{-1} + \hat{\mathbf{c}}_{j,2}\xi^{-2} + \hat{\mathbf{c}}_{j,3}\xi^{-3} + \ldots, \quad (j = 1, \ldots, q). \tag{5.5}$$

With A_j, C_j, and Q_j as in Section 4, it may then be verified that

$$Q_j A_j = \hat{A}_j Q_j, \text{ with } \hat{A}_j := \begin{bmatrix} -\mu_{j,d_j-1} & 1 & 0 & \cdots & 0 \\ -\mu_{j,d_j-2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\mu_{j,1} & 0 & 0 & \cdots & 1 \\ -\mu_{j,0} & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{5.6}$$

$$\text{and } C_j = \hat{C}_j Q_j, \text{ with } \hat{C}_j := \begin{bmatrix} \hat{\mathbf{c}}_{j,d_j} & \hat{\mathbf{c}}_{j,d_j-1} & \cdots & \hat{\mathbf{c}}_{j,1} \end{bmatrix}, \tag{5.7}$$

Finally, for $P_j \in \mathbb{R}^{d_j \times d_j}[\xi]$ as in (4.12), we define $\hat{D} \in \mathbb{R}^{m\times(n-m)}[\xi]$ as

$$\hat{D} := V_1 F_1^T + \sum_{j=1}^{q} \mathbf{v}_j \mathbf{a}_j^T + \mathbf{g}_j \mathbf{b}_j^T + \hat{C}_j P_j \hat{B}_j. \tag{5.8}$$

THEOREM 5.3. *The behavior $\mathcal{B}_{i/o}$ in (2.2) has the realization:*

$$\mathcal{B}_{i/o} = \left\{ \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix} \in \mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^n) \;\middle|\; \begin{array}{l} \exists \mathbf{x} \in \mathcal{L}_1^{loc}\left(\mathbb{R}, \mathbb{R}^d\right) \text{ with } \frac{d\mathbf{x}}{dt} = \hat{A}\mathbf{x} + \hat{B}\mathbf{u} \\ \qquad \text{and } \mathbf{y} = \hat{C}\mathbf{x} + \hat{D}\left(\frac{d}{dt}\right)\mathbf{u} \end{array} \right\},$$

where $\hat{A} = diag\begin{pmatrix} \hat{A}_1 & \cdots & \hat{A}_q \end{pmatrix}$, $\hat{B} = col\begin{pmatrix} \hat{B}_1 & \cdots & \hat{B}_q \end{pmatrix}$, $\hat{C} = \begin{bmatrix} \hat{C}_1 & \cdots & \hat{C}_q \end{bmatrix}$, and $\hat{D}, \hat{A}_j, \hat{B}_j$, and $\hat{C}_j$ are defined in equations (5.5) to (5.8) $(j = 1, \ldots, q)$. Moreover, $(\hat{C}, \hat{A})$ is observable, $(\hat{A}, \hat{B})$ is controllable if and only if $\mathcal{B}_{i/o}$ is behaviorally controllable, and $\Delta(\mathcal{B}_{i/o}) = \delta(\hat{D}) + d$ with $d = \sum_{j=1}^{q} d_j$.

Proof. Let A, B, C, and D be as in Theorem 4.1. Note initially that $\hat{D} - D = \sum_{j=1}^{q} \mathbf{g}_j(\mathbf{f}_j - \mathbf{b}_j)^T + (\mathbf{c}_j - \mathbf{v}_j)\mathbf{a}_j^T + \hat{C}_j P_j \hat{B}_j - C_j P_j B_j$ from (4.11) and (5.8). It may be verified that Q_j in (4.9) and P_j in (4.12) commute (the klth entry in P_jQ_j and Q_jP_j is equal to 0 when $l \leq k$, and $\sum_{i=l+1}^{k} \mu_{j,d_j-k+i}\xi^{i-l-1}$ otherwise, where $\mu_{j,d_j} := 1$). Thus, (4.7) and (5.7) imply $C_jP_jB_j = \hat{C}_jQ_jP_jB_j = \hat{C}_jP_jQ_jB_j = \hat{C}_jP_j\hat{B}_j$. Moreover, from (4.4), we obtain $\mathbf{g}_j(\mathbf{f}_j - \mathbf{b}_j)^T = \mathbf{g}_j\mu_j\mathbf{a}_j^T = (\mathbf{v}_j - \mathbf{c}_j)\mathbf{a}_j^T$. Hence, $\hat{D} = D$.

Now, let $Q := \text{diag}\begin{pmatrix} Q_1 & \cdots & Q_q \end{pmatrix}$, which is non-singular since $Q_1, \ldots, Q_q$ are non-singular. Then (4.7), (5.6), (5.7), and $\hat{D} = D$ imply

$$\begin{bmatrix} I_m & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} I_m & -D & -C \\ 0 & B & -\mathcal{A} \end{bmatrix} \begin{bmatrix} I_m & 0 & 0 \\ 0 & I_{n-m} & 0 \\ 0 & 0 & Q^{-1} \end{bmatrix} = \begin{bmatrix} I_m & -\hat{D} & -\hat{C} \\ 0 & \hat{B} & -\hat{\mathcal{A}} \end{bmatrix},$$

where $\hat{\mathcal{A}}(\xi) = \xi I_d - \hat{A}$. Thus, from (4.22), we obtain

$$\begin{bmatrix} I_m & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} W_1 & C \\ W_2 & \mathcal{A} \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & I_d \end{bmatrix} \begin{bmatrix} R_1 & -R_2 & 0 \\ X_1 & X_2 & -Q^{-1} \end{bmatrix} = \begin{bmatrix} I_m & -\hat{D} & -\hat{C} \\ 0 & \hat{B} & -\hat{\mathcal{A}} \end{bmatrix}.$$

Since Q is non-singular then the leftmost matrix in the above equation is unimodular. Further, from the proof of Theorem 4.1, the next two matrices in the above equation are also unimodular, and hence so too is the product of these three matrices. Furthermore, $\mathbf{x}$ is properly eliminable from $\mathcal{B}_s$ by Theorem 5.2, and so $\mathcal{B}_{i/o}$ satisfies (4.13) (see Section 2). That $(\hat{C}, \hat{A})$ is observable, $(\hat{A}, \hat{B})$ is controllable if and only if $\mathcal{B}_{i/o}$ is behaviorally controllable, and $\Delta(\mathcal{B}_{i/o}) = \delta(\hat{D}) + d$ then follows from a similar argument to the proof of Theorem 5.2. ∎

6. Efficient computations using companion matrices. In [15, 14, 6, 1], it is shown that the properties of companion matrices enable efficient computation. In this section, we extend the results in [15, 14] to construct explicit expressions for the trajectories of the realization in Theorem 4.1 of this paper, and we show how these can be computed efficiently (see Theorem 6.1). We also develop the results in [6, 1] to construct explicit expressions for the solutions to Lyapunov and Sylvester equations incorporating companion matrices (Theorems 6.2 and 6.4). This has relevance to model reduction, see e.g. [6]. In particular, we show how to efficiently compute the controllability and observability gramians for the realization in Theorem 4.1. We note that it is straightforward to obtain analogous results to the theorems in this section for the realization in Theorem 5.3.

Consider the realization in Theorem 4.1, and partition $\mathbf{x}$ compatibly with A as $\mathbf{x} =: \text{col}\begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_q \end{pmatrix}$. From the variation of the constants formula [17, Section 4.5], $\text{col}\begin{pmatrix} \mathbf{y} & \mathbf{u} \end{pmatrix} \in \mathcal{B}_{i/o}$ if and only if $\mathbf{y}(t) = D(\frac{d}{dt})\mathbf{u}(t) + \sum_{j=1}^{q} \int_{t=0}^{t=q} C_j e^{A_j(t-\tau)} B_j \mathbf{u}(\tau) d\tau + \sum_{j=1}^{q} C_j e^{A_j t} \mathbf{x}_j(0)$ for some $\mathbf{x}_j(0) \in \mathbb{R}^{\bullet}$ $(j = 1, \ldots, q)$. Here, $C_j e^{A_j t}$ and $C_j e^{A_j t} B_j$ can be computed efficiently using the following theorem, which extends results in [15, 14].

THEOREM 6.1. *Let A_j, B_j, and C_j be as in Theorem 4.1, where $\mathbf{c}_j$, $\mathbf{b}_j$, $\mathbf{Q}_j$ and $\mathbf{q}_j$ are as defined in Section 4. Further, let $\mathbf{z}_j := Q_j\mathbf{q}_j$, and let $M_j(\xi) =: (M_j)_{d_j-1}\xi^{d_j-1} + \ldots + (M_j)_1\xi + (M_j)_0$ (resp. $N_j(\xi) =: (N_j)_{d_j-1}\xi^{d_j-1} + \ldots + (N_j)_1\xi + (N_j)_0$), where M_j (resp. N_j) is the remainder on division of $\mathbf{c}_j\mathbf{z}_j^T$ (resp. $\mathbf{c}_j\mathbf{b}_j^T$) by μ_j. Then*

$$C_j e^{A_j t} = \sum_{k=1}^{d_j} (M_j)_{k-1}(\Phi_j)_{k,d_j}(t), \;\; \textit{and} \;\; C_j e^{A_j t} B = \sum_{k=1}^{d_j} (N_j)_{k-1}(\Phi_j)_{k,d_j}(t), \tag{6.1}$$

where $(\Phi_j)_{k,d_j}$ *denotes the element in the* k*th row and* d_j*th column of* Φ_j *with* $\Phi_j(t) := e^{A_j t}$ $(t \in \mathbb{R})$*. This has the power series* $(\Phi_j)_{k,d_j}(t) = \sum_{l=0}^{\infty} h_{l+k} t^l / l!$ $(k = 1, \ldots, d_j)$*, where* $h_1 = \ldots = h_{d_j-1} = 0$*,* $h_{d_j} = 1$*, and* $h_{l+1} = -\sum_{i=1}^{d_j} h_{l+1-i} \mu_{j,d_j-i}$ $(l = d_j, d_j+1, \ldots)$*.*

As a consequence of the above theorem, it is not necessary to compute e^{At} to obtain the trajectories of the realization in Theorem 4.1. Instead, these may be efficiently computed from the entries $(\Phi_j)_{k,d_j}$, which may be approximated numerically using the power series provided in that theorem (we refer also to [14] for a more detailed discussion on the efficient numerical approximation of $(\Phi_j)_{k,d_j}$).

To show Theorem 6.1, we first let $(\phi_j)_k$ denote the kth column of the identity matrix I_{d_j} $(k = 1, \ldots, d_j)$, and we let $(\phi_j)_0 = 0$. Then from (4.10) we find

$$(\phi_j)_k = A_j (\phi_j)_{k+1} + \mu_{j,k} (\phi_j)_{d_j} \text{ for } k = 0, \ldots, d_j - 1. \tag{6.2}$$

Since, in addition $e^{A_j t} = \sum_{l=0}^{\infty} A_j^l t^l / l!$ commutes with A_j, then the above equation also holds when $(\phi_j)_k$ denotes the kth column of $\Phi_j(t) := e^{A_j t}$ (note that this is also the case for $\Phi_j = (sI - A_j)^{-1}$). It follows that the entries $(\Phi_j)_{k,l}$ in Φ_j may be routinely computed from the entries in the final column of Φ_j using the recursion:

$$\begin{aligned}
&(\Phi_j)_{k,l} = (\Phi_j)_{k+1,l+1} + \mu_{j,l} (\Phi_j)_{k,d_j}, && l = d_j - 1, \ldots, 1,\ k = 1, \ldots, l,\\
&(\Phi_j)_{k,1} = -\mu_{j,0} (\Phi_j)_{k-1,d_j}, && k = 2, \ldots, d_j,\\
&(\Phi_j)_{k,l} = (\Phi_j)_{k-1,l-1} - \mu_{j,l} (\Phi_j)_{k,d_j}, && l = 2, \ldots, d_j - 1,\ k = l+1, \ldots, d_j.
\end{aligned}$$

Moreover, from [15], $(\Phi_j)_{k,d_j}$ is the sum of the residues of $(\xi^{k-1} e^{\xi t}) / \mu_j(\xi)$ $(k = 1, \ldots, d_j)$. In particular, from the power series $\Phi_j(t) = \sum_{l=0}^{\infty} A_j^l t^l / l!$, we obtain the power series in Theorem 6.1.

Let Q_j and $\mathbf{e}_j$ be as in (4.9) and (4.16), respectively (so $\Phi_j \mathbf{e}_j$ is the final column in Φ_j). Firstly, note that (6.2) implies $\Phi_j = \begin{bmatrix} A_j^{d_j-1} \Phi_j \mathbf{e}_j & A_j^{d_j-2} \Phi_j \mathbf{e}_j & \cdots & \Phi_j \mathbf{e}_j \end{bmatrix} Q_j$. Secondly, it follows from the preceding paragraph that $\Phi_j \mathbf{e}_j$ is the sum of the residues of $(\mathbf{p}_j(\xi) e^{\xi t}) / \mu_j(\xi)$ where $\mathbf{p}_j$ is as in (4.14), whence from (4.10) it is evident that $A_j^l \Phi_j \mathbf{e}_j$ is the sum of the residues of $(\mathbf{p}_j(\xi) \xi^l e^{\xi t}) / \mu_j(\xi)$ $(l = 1, 2, \ldots)$. Then, with $\mathbf{q}_j$ is as in (4.15), we conclude that $\begin{bmatrix} A_j^{d_j-1} \Phi_j \mathbf{e}_j & A_j^{d_j-2} \Phi_j \mathbf{e}_j & \cdots & \Phi_j \mathbf{e}_j \end{bmatrix}$ is the sum of the residues of $(\mathbf{p}_j(\xi) \mathbf{q}_j(\xi)^T e^{\xi t}) / \mu_j(\xi)$. It follows from (4.24) that $C_j e^{A_j t}$ is the sum of the residues of $C_j \mathbf{p}_j(\xi) \mathbf{q}_j(\xi)^T Q_j e^{\xi t} / \mu_j(\xi) = \mathbf{c}_j(\xi) \mathbf{z}_j(\xi)^T e^{\xi t} / \mu_j(\xi)$, where $\mathbf{z}_j := Q_j \mathbf{q}_j$. Moreover, from (4.7) and (4.24), we find that $C_j e^{A_j t} B$ is the sum of the residues of $\mathbf{c}_j(\xi) \mathbf{b}_j(\xi)^T e^{\xi t} / \mu_j(\xi)$. Theorem 6.1 then follows since the sum of the residues of $\mathbf{c}_j(\xi) \mathbf{z}_j(\xi)^T e^{\xi t} / \mu_j(\xi)$ (resp. $\mathbf{c}_j(\xi) \mathbf{b}_j(\xi)^T e^{\xi t} / \mu_j(\xi)$) is equal to the sum of the residues of $M_j(\xi) e^{\xi t} / \mu_j(\xi)$ (resp. $N_j(\xi) e^{\xi t} / \mu_j(\xi)$).

We now apply Theorem 6.1 to the example following Theorem 4.1. Here, $\mu_1(\xi) = \xi^2 + 2\xi + 1$, so $\mathbf{z}_1(\xi) = \operatorname{col}\begin{pmatrix} \xi + 2 & 1 \end{pmatrix}$, and dividing $\mathbf{c}_1 \mathbf{z}_1^T$ and $\mathbf{c}_1 \mathbf{b}_1^T$ by μ_1 gives

$$M(\xi) = \begin{bmatrix} 2\xi+2 & 2\xi+2 \\ 1 & -\xi \\ \xi+2 & 1 \end{bmatrix}, \text{ and } N(\xi) = \begin{bmatrix} 2\xi+2 & -2\xi-2 \\ -3\xi-2 & \xi \\ -2\xi-1 & -1 \end{bmatrix}. \tag{6.3}$$

Next, since $d_1 = 2$, we must compute the entries $(\Phi_1)_{1,2}$ and $(\Phi_1)_{2,2}$, which are equal to the residues of $e^{\xi t} / (\xi + 1)^2$ and $(\xi e^{\xi t}) / (\xi + 1)^2$ at $\xi = -1$, respectively. These may be evaluated directly in this case as the roots of ξ are rational. Alternatively, in the power series in Theorem 6.1, h_l $(l \geq 1)$ is the solution to the difference equation $h_{l+1} + 2h_l + h_{l-1} = 0$ with $h_1 = 0$ and $h_2 = 1$, so $h_l = (l-1)(-1)^l$. We then obtain

$(\Phi_1)_{1,2}(t) = t\sum_{l=0}^{\infty}(-t)^l/l! = te^{-t}$, and $(\Phi_1)_{2,2}(t) = \sum_{l=0}^{\infty}(-t)^l/l! - t\sum_{l=0}^{\infty}(-t)^l/l! = (1-t)e^{-t}$. By combining these expressions with (6.3), we find

$$C_1 e^{A_1 t} = \begin{bmatrix} 2 & 2 \\ 0 & -1 \\ 1 & 0 \end{bmatrix} e^{-t} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} te^{-t}, \text{ and } C_1 e^{A_1 t} B_1 = \begin{bmatrix} 2 & -2 \\ -3 & 1 \\ -2 & 0 \end{bmatrix} e^{-t} + \begin{bmatrix} 0 & 0 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} te^{-t}.$$

In the remainder of this section, we extend results from [6, 1] on the efficient computation of solutions to Lyapunov and Sylvester equations involving companion matrices. This has relevance to model reduction [6]. Here, we show how this may be used for the efficient computation of the symmetric solutions for X and $\hat{X}$ to the Lyapunov equations $A^T X + XA = -Z$ and $\hat{X}A^T + A\hat{X} = -\hat{Z}$ (with $Z, \hat{Z}$ symmetric) for the realization in Theorem 4.1. In particular, we show how to compute the observability and controllability gramians for the system in the preceding example.

By partitioning X, $\hat{X}$, Z and $\hat{Z}$ compatibly with A as:

$$X = \begin{bmatrix} X_{1,1} & \cdots & X_{1,q} \\ \vdots & \ddots & \vdots \\ X_{q,1} & \cdots & X_{q,q} \end{bmatrix}, \hat{X} = \begin{bmatrix} \hat{X}_{1,1} & \cdots & \hat{X}_{1,q} \\ \vdots & \ddots & \vdots \\ \hat{X}_{q,1} & \cdots & \hat{X}_{q,q} \end{bmatrix}, Z = \begin{bmatrix} Z_{1,1} & \cdots & Z_{1,q} \\ \vdots & \ddots & \vdots \\ Z_{q,1} & \cdots & Z_{q,q} \end{bmatrix}, \text{and } \hat{Z} = \begin{bmatrix} \hat{Z}_{1,1} & \cdots & \hat{Z}_{1,q} \\ \vdots & \ddots & \vdots \\ \hat{Z}_{q,1} & \cdots & \hat{Z}_{q,q} \end{bmatrix}, \tag{6.4}$$

then we seek solutions for $X_{i,j}$ and $\hat{X}_{i,j}$ to $A_i^T X_{i,j} + X_{i,j}A_j = -Z_{i,j}$ and $\hat{X}_{i,j}A_j^T + A_i\hat{X}_{i,j} = -\hat{Z}_{i,j}$. Moreover, given the symmetry of $X, \hat{X}, Z$ and $\hat{Z}$, we need only consider the case $i \leq j$. Here, $\hat{X}_{i,j}$ may be efficiently computed using Theorem 6.4, and $X_{i,j}$ using the following theorem:

THEOREM 6.2. *Let A_j, μ_j, and $\mathbf{p}_j$ be as in Section 4 ($j = 1, \ldots, q$). Now, let i, j be integers with $1 \leq i \leq j \leq q$, and let $Z_{i,j} \in \mathbb{R}^{d_i \times d_j}$. If there exists a solution $X_{i,j} \in \mathbb{R}^{d_i \times d_j}$ to $A_i^T X_{i,j} + X_{i,j}A_j = -Z_{i,j}$, then there exist $u_{i,j}, v_{i,j} \in \mathbb{R}[\xi]$ with*

$$\mu_i(-\xi)u_{i,j}(\xi) + \mu_j(\xi)v_{i,j}(-\xi) = z_{i,j}(-\xi, \xi), \text{ where } z_{i,j}(\eta, \xi) = \mathbf{p}_i(\eta)^T Z_{i,j}\mathbf{p}_j(\xi). \tag{6.5}$$

Furthermore, $u_{i,j}$ and $v_{i,j}$ may be chosen such that their degrees are less than d_i and d_j, respectively, in which case a solution $X_{i,j} \in \mathbb{R}^{d_i \times d_j}$ to $A_i^T X_{i,j} + X_{i,j}A_j = -Z_{i,j}$ is obtained by equating coefficients in the equation:

$$(\eta + \xi)\mathbf{p}_i(\eta)^T X_{i,j}\mathbf{p}_j(\xi) = \mu_i(\eta)u_{i,j}(\xi) + \mu_j(\xi)v_{i,j}(\eta) - z_{i,j}(\eta, \xi). \tag{6.6}$$

In particular, $X_{i,j}$ in (6.6) can be obtained by the recursive equations:

$$\begin{aligned}
(X_{i,j})_{d_i,k} &= (u_{i,j})_{k-1}, \quad k = 1, \ldots, d_j, \\
(X_{i,j})_{k,d_j} &= (v_{i,j})_{k-1}, \quad k = 1, \ldots, d_i - 1, \\
(X_{i,j})_{l,k} &= \mu_{i,l}(X_{i,j})_{d_i,k} + \mu_{j,k-1}(X_{i,j})_{l+1,d_j} - (Z_{i,j})_{l+1,k} - (X_{i,j})_{l+1,k-1} \\
&\qquad \textit{with } (X_{i,j})_{l+1,0} := 0, \qquad l = d_i - 1, \ldots, 1, \qquad k = 1, \ldots, d_j - d_i + l, \\
(X_{i,j})_{k,l} &= \mu_{i,k-1}(X_{i,j})_{d_i,l+1} + \mu_{j,l}(X_{i,j})_{k,d_j} - (Z_{i,j})_{k,l+1} - (X_{i,j})_{k-1,l+1} \\
&\qquad \textit{with } (X_{i,j})_{0,l+1} := 0, \quad l = d_j - 1, \ldots, d_j - d_i + 2, \quad k = 1, \ldots, d_i - d_j + l - 1,
\end{aligned}$$

where $(Z_{i,j})_{k,l}$ (resp. $(X_{i,j})_{k,l}$) denotes the entry in the kth column and lth row of $Z_{i,j}$ (resp $X_{i,j}$), and $(v_{i,j})_{k-1}$ (resp. $(u_{i,j})_{l-1}$) denotes the coefficient of ξ^{k-1} (resp. ξ^{l-1}) in $v_{i,j}$ (resp. $u_{i,j}$), for $k = 1, \ldots, d_i$, $l = 1, \ldots, d_j$.

To see Theorem 6.2, we note initially that the first (resp. second, third, fourth) of the recursive equations follows by equating coefficients of $\eta^{d_i}\xi^{k-1}$ (resp. $\eta^{k-1}\xi^{d_j}$, $\eta^l\xi^{k-1}$, $\eta^{k-1}\xi^l$) in (6.6). Now, consider the $d_i \times d_i$ matrices:

$$J_i := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (-1)^{d_i} \end{bmatrix}, \text{ and } K_i := \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}, \; i = 1, \ldots, q. \tag{6.7}$$

Note that $J_i(A_i^T X + XA_j + Z) = (J_i A_i^T J_i)(J_i X) + (J_i X)A_j + J_i Z$, and $J_i A_i^T J_i$ is the negative of the transpose of the companion matrix for $\mu_i(-\xi)$. Theorem 6.2 may then be shown using [6, Theorem 1].

The polynomials $u_{i,j}, v_{i,j}$ in (6.5) can be computed efficiently using algorithms in symbolic algebra programs. These amount to solving an equation of the form $\mathcal{S}^T \mathrm{col}\begin{pmatrix}\mathbf{u}_{i,j} & \mathbf{v}_{i,j}\end{pmatrix} = \tilde{\mathbf{z}}_{i,j}$, in which $\mathcal{S}$ is a Sylvester matrix for $\mu_i(-\xi)$ and $\mu_j(\xi)$, $\tilde{\mathbf{z}}_{\mathbf{i},\mathbf{j}}$ is a vector of coefficients of $\mathbf{z}(-\xi,\xi)$, and $\mathbf{u}_{i,j}$, $\mathbf{v}_{i,j}$ are vectors of coefficients of $u_{i,j}(\xi)$ and $v_{i,j}(-\xi)$, respectively. This yields an algorithm for the computation of the solution to $A_i^T X_{i,j} + X_{i,j}A_j = -Z_{i,j}$ in order d_j^2 arithmetic operations.

Theorem 6.2 generalizes [1, Theorem 1], which considers the Lyapunov equation $A_i^T X_{i,i} + X_{i,i}A_i = -Z_{i,i}$. The approach in [1] did not explicitly invoke the polynomial equation (6.5). To recover the results in [1], we note that since $X_{i,i}$ is symmetric then we need only evaluate the first and the third of the recursive relationships in Theorem 6.2, and we require $v_{i,i}(-\xi) = u_{i,i}(-\xi)$ in (6.5). Furthermore, $z_{i,i}(-\xi,\xi)$ in (6.5) is an even polynomial. Now, consider decomposing μ_i and u_i into even and odd parts ($\mu_i =: \mu_i^{(e)} + \mu_i^{(o)}$ and $u_i =: u_i^{(e)} + u_i^{(o)}$). Here, $\mu_i^{(e)}(\xi) = \mu_{i,0} + \mu_{i,2}\xi^2 + \ldots$; $\mu_i^{(o)} = \mu_{i,1}\xi + \mu_{i,3}\xi^3 + \ldots$; $u_i^{(e)}$ and $u_i^{(o)}$ are defined analogously; and, since $\mu_j(\xi) = \mu_i(\xi)$ and $v_{i,i}(-\xi) = u_{i,i}(-\xi)$, then (6.5) implies $\mu_i^{(e)}(\xi)u_i^{(e)}(\xi) - \mu_i^{(o)}(\xi)u_i^{(o)}(\xi) = (1/2)z_{i,i}(-\xi,\xi)$. This can be solved for $u_i^{(e)}$ and $u_i^{(o)}$ by considering a matrix equation involving the transpose of a Sylvester matrix for $\mu_i^{(e)}(\xi)$ and $\mu_i^{(o)}(\xi)$ (c.f. [1, Equation (8)]).

Now, consider again the example following Theorem 4.1. Suppose we want to solve $A^T X + XA = -C^T C$ to obtain the observability gramian for this system. In this case, X takes the form of (6.4) with $q = 2$, $X_{2,1} = X_{1,2}^T$, and where $X_{1,1}, X_{1,2}$, and $X_{2,2}$ can be obtained from Theorem 6.2. Specifically, for $X_{2,2}$, we have $Z_{2,2} = C_2^T C_2$, whence from (4.24) and (6.5) we obtain $z_{2,2}(\eta,\xi) = \mathbf{p}_2(\eta)^T C_2^T C_2 \mathbf{p}_2(\xi) = \mathbf{c}_2(\eta)^T \mathbf{c}_2(\xi) = (3\eta^2 + 6\eta + 2)(3\xi^2 + 6\xi + 2) + (1+\eta)(1+\xi) + 1$. Using the Maple command `gcdex`, we obtain $u_{2,2}(\xi) = 41\xi^2/18 + 83\xi/18 + 3/2 = v_{2,2}(\xi)$, which gives the entries in the last row and column of $X_{2,2}$. The recursive relationships in Theorem 6.2 then give

$$(6.8)\quad X_{2,2} = \begin{bmatrix} \frac{67}{18} & \frac{41}{9} & \frac{3}{2} \\ \frac{41}{9} & \frac{31}{3} & \frac{83}{18} \\ \frac{3}{2} & \frac{83}{18} & \frac{41}{18} \end{bmatrix}, \text{ which satisfies } A_2^T X_{2,2} + X_{2,2}A_2 = -C_2^T C_2 = -\begin{bmatrix} 6 & 13 & 6 \\ 13 & 37 & 18 \\ 6 & 18 & 9 \end{bmatrix}.$$

We now consider the equations $\hat{X}_{i,j}A_j^T + A_i\hat{X}_{i,j} = -\hat{Z}_{i,j}$. These can be solved efficiently using Bezoutians and applying results in [2]. We first show a lemma.

LEMMA 6.3. *Let $\mu_i, \mu_j \in \mathbb{R}[\xi]$ have degrees d_i and d_j, respectively, with $d_j \geq d_i$. There is a unique matrix $\mathcal{B}(\mu_j(\xi), \mu_i(\xi)) \in \mathbb{R}^{d_j \times d_j}$ which satisfies $\mu_j(\eta)\mu_i(\xi) - \mu_j(\xi)\mu_i(\eta) = (\eta - \xi)\mathbf{p}_j(\eta)^T \mathcal{B}(\mu_j(\xi), \mu_i(\xi))\mathbf{p}_j(\xi)$, where $\mathbf{p}_j$ is as in (4.14). Now, let $r \in \mathbb{R}[\xi]$ with $r(\xi) =: r_{d_j-1}\xi^{d_j-1} + \ldots + r_1\xi + r_0$. If there exist $u, w \in \mathbb{R}[\xi]$ satisfying $\mu_i(\xi)u(\xi) + \mu_j(\xi)w(\xi) = r(\xi)$, then there exists col$\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_j}\end{pmatrix} \in \mathbb{R}^{d_j}$ satisfying*

$$(6.9)\qquad \mathcal{B}(\mu_j(\xi), \mu_i(\xi))\,col\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_j}\end{pmatrix} = col\begin{pmatrix}r_0 & \ldots & r_{d_j-1}\end{pmatrix}.$$

Furthermore, whenever col$\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_j}\end{pmatrix} \in \mathbb{R}^{d_j}$ satisfies (6.9), then there exists $u, w \in \mathbb{R}[\xi]$ with $\mu_i(\xi)u(\xi) + \mu_j(\xi)w(\xi) = r(\xi)$ where $u(\xi)/\mu_j(\xi)$ has the formal series expansion $u(\xi)/\mu_j(\xi) = \hat{u}_1/\xi + \hat{u}_2/\xi^2 + \ldots$, with $\hat{u}_k = -\sum_{l=1}^{d_j} \hat{u}_{k-l}\mu_{j,d_j-l}$ for $k > d_j$.

To see this lemma, note initially that $\mathcal{B}(\mu_j(\xi), \mu_i(\xi))$ is uniquely defined by (6.9) since $(\eta - \xi)$ divides $\nu_2(\eta)\nu_1(\xi) - \nu_2(\xi)\nu_1(\eta)$ by the factor theorem (this matrix is known as the Bezoutian of μ_j and μ_i). We then note that if $\mu_i(\xi)u(\xi) + \mu_j(\xi)w(\xi) = r(\xi)$ where $\deg(r(\xi)) < d_j$, and $u(\xi)/\mu_j(\xi)$ has the formal series expansion $u(\xi)/\mu_j(\xi) =$

$\hat{u}_1/\xi+\hat{u}_2/\xi^2+\ldots$, then in the formal series expansion $w(\xi)/\mu_i(\xi)=\hat{w}_1/\xi+\hat{w}_2/\xi^2+\ldots$ we require $\hat{w}_k=-\hat{u}_k$ $(k=1,\ldots,d_i)$. Thus, with $Q_j, \mathbf{p}_j(\xi)$, and K_j as in (4.9), (4.14), and (6.7), respectively, it follows that $u(\xi)=\mathbf{p}_j(\xi)^T K_j Q_j \text{col}\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_j}\end{pmatrix}$ and $w(\xi)=-\mathbf{p}_i(\xi)^T K_i Q_i \text{col}\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_i}\end{pmatrix}$ (c.f. equations (4.5) to (4.7)). Then, by equating coefficients in $u(\xi)\mu_i(\xi)+w(\xi)\mu_j(\xi)=r(\xi)$, noting that $\mathcal{B}(\mu_j(\xi),\mu_i(\xi))$ is symmetric and has the representation in [2, equation (1.1)], we obtain (6.9). Moreover, any solution $\text{col}\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_j}\end{pmatrix}$ to (6.9) defines polynomials $u(\xi):=\mathbf{p}_j(\xi)^T K_j Q_j \text{col}\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_j}\end{pmatrix}$ and $w(\xi):=-\mathbf{p}_i(\xi)^T K_i Q_i \text{col}\begin{pmatrix}\hat{u}_1 & \cdots & \hat{u}_{d_i}\end{pmatrix}$ which satisfy $\mu_i(\xi)u(\xi)+\mu_j(\xi)w(\xi)=r(\xi)$, where $u(\xi)/\mu_j(\xi)$ has the formal series expansion indicated in Lemma 6.3.

We will now show Theorem 6.4. Note that equation (6.10) can be efficiently solved using the methods in [2], whence Theorem 6.4 provides an algorithm to find solutions $\hat{X}_{i,j}$ to the equation $\hat{X}_{i,j}A_j^T+A_i\hat{X}_{i,j}=-\hat{Z}_{i,j}$ in order d_j^2 arithmetic operations.

THEOREM 6.4. *Let A_j,μ_j,Q_j be as in Section 4, and let K_i be as in (6.7) $(i,j=1,\ldots,q)$. Now, let i,j be integers satisfying $1\le i\le j\le q$, let $\hat{Z}_{i,j}\in\mathbb{R}^{d_i\times d_j}$, and let $z_{i,j}(\eta,\xi)=\mathbf{p}_i(\eta)^T K_i Q_i \hat{Z}_{i,j}(K_j Q_j)^T\mathbf{p}_j(\xi)$. Finally, let $r_{i,j}(\xi)=:(r_{i,j})_{d_j-1}\xi^{d_j-1}+\ldots+(r_{i,j})_1\xi+(r_{i,j})_0$ be the remainder in the division of $z_{i,j}(-\xi,\xi)$ by $\mu_j(\xi)$, and let $\mathcal{B}(\mu_j(\xi),\mu_i(-\xi))$ satisfy $\mu_j(\eta)\mu_i(-\xi)-\mu_j(\xi)\mu_i(-\eta)=(\eta-\xi)\mathbf{p}_j(\eta)^T\mathcal{B}(\mu_j(\xi),\mu_i(-\xi))\mathbf{p}_j(\xi)$. If there exists a solution $\hat{X}_{i,j}\in\mathbb{R}^{d_i\times d_j}$ to $\hat{X}_{i,j}A_j^T+A_i\hat{X}_{i,j}=-\hat{Z}_{i,j}$, then there exists a solution $\text{col}\begin{pmatrix}(\hat{u}_{i,j})_1 & \cdots & (\hat{u}_{i,j})_{d_j}\end{pmatrix}\in\mathbb{R}^{d_j}$ to*

$$\mathcal{B}(\mu_j(\xi),\mu_i(-\xi))\text{col}\begin{pmatrix}(\hat{u}_{i,j})_1 & \cdots & (\hat{u}_{i,j})_{d_j}\end{pmatrix}=\text{col}\begin{pmatrix}(r_{i,j})_0 & \ldots & (r_{i,j})_{d_j-1}\end{pmatrix}. \tag{6.10}$$

Furthermore, if $\text{col}\begin{pmatrix}(\hat{u}_{i,j})_1 & \cdots & (\hat{u}_{i,j})_{d_j}\end{pmatrix}\in\mathbb{R}^{d_j}$ satisfies (6.10), then there is a solution $\hat{X}_{i,j}$ to $\hat{X}_{i,j}A_i^T+A_j\hat{X}=-\hat{Z}_{i,j}$ for which the entries $(X_{i,j})_{k,l}$ in the kth column and lth row of $X_{i,j}$ $(k=1,\ldots,d_i, l=1,\ldots,d_j)$ satisfy

$$(\hat{X}_{i,j})_{1,l}:=(\hat{u}_{i,j})_l,\quad l=1,\ldots,d_j,$$
$$\textit{and } (\hat{X}_{i,j})_{k,l}:=-(\hat{Z}_{i,j})_{k-1,l}-(\hat{X}_{i,j})_{k-1,l+1},\quad l=1,\ldots,d_j, k=2,\ldots,d_i,$$
$$\textit{with } (\hat{X}_{i,j})_{k-1,d_j+1}:=-\sum_{l=1}^{d_j}(\hat{X}_{i,j})_{k-1,l}\mu_{j,l-1},\quad k=2,\ldots,d_i.$$

To show the above theorem, we note initially that from (4.10) and (5.6) it is evident that $\hat{A}_j=K_jA_j^TK_j$ and so $K_jQ_jA_j=A_j^TK_jQ_j$ $(j=1,\ldots,q)$. Since K_jQ_j is symmetric, we conclude that $\hat{X}_{i,j}$ solves $\hat{X}_{i,j}A_j^T+A_i\hat{X}_{i,j}=-\hat{Z}_{i,j}$ if and only if $X_{i,j}:=K_iQ_i\hat{X}_{i,j}(K_jQ_j)^T$ solves $X_{i,j}A_j+A_i^TX_{i,j}=-Z_{i,j}$, where $Z_{i,j}:=K_iQ_i\hat{Z}_{i,j}(K_jQ_j)^T$. It follows from Theorem 6.2 that whenever this has a solution then there exist $u_{i,j},v_{i,j}\in\mathbb{R}[\xi]$ satisfying (6.5). Now, let $z_{i,j}(-\xi,\xi)=q_{i,j}(\xi)\mu_j(\xi)+r_{i,j}(\xi)$ (i.e. $q_{i,j}(\xi)$ is the quotient on division of $z_{i,j}(-\xi,\xi)$ by $\mu_j(\xi)$), and note that (6.5) implies $u_{i,j}(\xi)\mu_i(-\xi)+w_{i,j}(\xi)\mu_j(\xi)=r_{i,j}(\xi)$, where $w_{i,j}(\xi):=v_{i,j}(-\xi)-q_{i,j}(\xi)$. Then, by substituting $\mu_i(-\xi)$ for $\mu_i(\xi)$ in Lemma 6.3, we conclude that if there exists an $\hat{X}_{i,j}\in\mathbb{R}^{d_i\times d_j}$ satisfying $\hat{X}_{i,j}A_i^T+A_j\hat{X}=-\hat{Z}_{i,j}$ then there exists a solution $\text{col}\begin{pmatrix}(\hat{u}_{i,j})_1 & \cdots & (\hat{u}_{i,j})_{d_j}\end{pmatrix}\in\mathbb{R}^{d_j}$ to (6.10), and if $\text{col}\begin{pmatrix}(\hat{u}_{i,j})_1 & \cdots & (\hat{u}_{i,j})_{d_j}\end{pmatrix}\in\mathbb{R}^{d_j}$ satsfies (6.10) then there is a solution to $u_{i,j}(\xi)\mu_i(-\xi)+w_{i,j}(\xi)\mu_j(\xi)=r_{i,j}(\xi)$ for which $u(\xi)/\mu_j(\xi)$ has the formal series expansion $u(\xi)/\mu_j(\xi)=(\hat{u}_{i,j})_1/\xi+(\hat{u}_{i,j})_2/\xi^2+\ldots$.

It remains to show the recursive equations in Theorem 6.4. To show these, we first consider an arbitrary matrix $\Omega\in\mathbb{R}^{d_i\times d_j}$ and the associated two variable polynomial $\omega(\eta,\xi):=\mathbf{p}_i(\eta)^T\Omega\mathbf{p}_j(\xi)$. We note that $\omega(\eta,\xi)/(\mu_i(\eta)\mu_j(\xi))$ has a two variable formal series expansion $\omega(\eta,\xi)/(\mu_i(\eta)\mu_j(\xi)) = \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\hat{\omega}_{k,l}\eta^{-k}\xi^{-l}$. Furthermore, the matrix $\hat{\Omega}$ with entries $\hat{\Omega}_{k,l} := \hat{\omega}_{k,l}$ $(k = 1,\ldots,d_i, l = 1,\ldots,d_j)$ satisfies $\Omega = K_iQ_i\hat{\Omega}(K_jQ_j)^T$. Accordingly, suppose there exist $u_{i,j}, v_{i,j} \in \mathbb{R}[\xi]$ satisfying (6.6); let $x_{i,j}(\eta,\xi) = \mathbf{p}_i(\eta)^T X_{i,j}\mathbf{p}_j(\xi)$ satisfy (6.6); and consider the formal series expansions $x_{i,j}(\eta,\xi)/(\mu_i(\eta)\mu_j(\xi)) = \sum_{k,l=1}^{\infty}(\hat{x}_{i,j})_{k,l}\eta^{-k}\xi^{-l}$, $z_{i,j}(\eta,\xi)/(\mu_i(\eta)\mu_j(\xi)) = \sum_{k,l=1}^{\infty}(\hat{z}_{i,j})_{k,l}\eta^{-k}\xi^{-l}$, $u_{i,j}(\xi)/\mu_j(\xi) = \sum_{k=1}^{\infty}(\hat{u}_{i,j})_k\xi^{-k}$, $v_{i,j}(\eta)/\mu_i(\eta) = \sum_{k=1}^{\infty}(\hat{v}_{i,j})_k\eta^{-k}$. Then (6.6) implies

$$(\eta+\xi)\sum_{k,l=1}^{\infty}\frac{(\hat{x}_{i,j})_{k,l}}{\eta^k\xi^l} = \sum_{k=1}^{\infty}\frac{(\hat{u}_{i,j})_k}{\xi^k} + \sum_{k=1}^{\infty}\frac{(\hat{v}_{i,j})_k}{\eta^k} - \sum_{k,l=1}^{\infty}\frac{(\hat{z}_{i,j})_{k,l}}{\eta^k\xi^l}. \tag{6.11}$$

Since, by Theorem 6.2, there is a solution $X_{i,j}$ to $X_{i,j}A_j + A_i^T X_{i,j} = -Z_{i,j}$ which satisfies (6.6), then it follows that there is a solution $\hat{X}_{i,j}$ to $\hat{X}_{i,j}A_j^T + A_i\hat{X}_{i,j} = -\hat{Z}_{i,j}$ whose entries $(\hat{X}_{i,j})_{k,l}$ satisfy $(\hat{X}_{i,j})_{k,l} = (\hat{x}_{i,j})_{k,l}$ $(k = 1,\ldots,d_i, l = 1,\ldots,d_j)$. The first (resp. second) of the recursive equations in Theorem 6.4 then comes from equating coefficients of ξ^{-l} (resp. $\eta^{1-k}\xi^{-l}$) in (6.11), and for the third equation we note that equating coefficients of $\eta^{1-k}\xi^{-1}$ in $x(\eta,\xi)/\mu_i(\eta) = \sum_{k,l=1}^{\infty}(\hat{x}_{i,j})_{k,l}\eta^{-k}\xi^{-l}\mu_j(\xi)$ gives $(\hat{x}_{i,j})_{k-1,d_j+1} = -\sum_{l=1}^{d_j}(\hat{x}_{i,j})_{k-1,l}\mu_{j,l-1}$.

Finally in this section, we consider again the example following Theorem 4.1, and we show how to solve $\hat{X}A^T + A\hat{X} = -BB^T$ to obtain the controllability gramian for this system. As before, $q = 2$ and $\hat{X}_{2,1} = \hat{X}_{1,2}^T$ in (6.4), and in this case we will find $\hat{X}_{1,1}, \hat{X}_{1,2}$, and $\hat{X}_{2,2}$ from Theorem 6.4. For $\hat{X}_{1,2}$, we have $\hat{Z}_{1,2} = B_1B_2^T$, and we note from (4.14), (4.15) and (6.7) that $\mathbf{q}_j = K_j\mathbf{p}_j$, so from (4.7) and (4.24) we obtain $z_{1,2}(\eta,\xi) = \mathbf{q}_1(\eta)Q_1\hat{B}_1B_2Q_2\mathbf{q}_2(\xi) = \mathbf{b}_1(\eta)^T\mathbf{b}_2(\xi) = (-2\eta-1)-\xi$. We then obtain $r_{1,2}(\xi)$ as the remainder in the division of $z_{1,2}(-\xi,\xi)$ by $\mu_2(\xi)$ which in this case is equal to $z_{1,2}(-\xi,\xi)$ itself. Then in (6.10) we have

$$\begin{bmatrix} 9 & 2 & 1 \\ 2 & -12 & -2 \\ 1 & -2 & 1 \end{bmatrix}\begin{bmatrix} (\hat{u}_{1,2})_1 \\ (\hat{u}_{1,2})_2 \\ (\hat{u}_{1,2})_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \text{ where } \begin{bmatrix} 1 & \eta & \eta^2 \end{bmatrix}\begin{bmatrix} 9 & 2 & 1 \\ 2 & -12 & -2 \\ 1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ \xi \\ \xi^2 \end{bmatrix} = \frac{\mu_2(\eta)\mu_1(-\xi)-\mu_2(\xi)\mu_1(-\eta)}{\eta-\xi}.$$

By solving the above, we obtain $(\hat{u}_{1,2})_1 = (\hat{u}_{1,2})_2 = (\hat{u}_{1,2})_3 = -1/12$, which gives the entries in the first row of $\hat{X}_{1,2}$. Finally, the recursive equations in Theorem 6.4 give

$$\hat{X}_{1,2} = \begin{bmatrix} -\frac{1}{12} & -\frac{1}{12} & -\frac{1}{12} \\ \frac{1}{12} & \frac{1}{12} & \frac{13}{12} \end{bmatrix}, \text{ which satisfies } \hat{X}_{1,2}A_2^T + A_1\hat{X}_{1,2} = -B_1B_2^T = -\begin{bmatrix} 0 & 0 & -2 \\ 0 & -1 & 7 \end{bmatrix}.$$

7. Input-output representations for behaviors. In this section, we show how the results in this paper may be used to construct a realization for a general behavior $\mathcal{B}$ as in (2.1). We will first show the following lemma:

LEMMA 7.1. *Let $\mathcal{B}_{i/o}$ be as in (2.2), and let $T \in \mathbb{R}^{n\times n}$ be non-singular. Then*

$$\mathcal{B} = \left\{ T\,col\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \,\middle|\, col\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \in \mathcal{B}_{i/o} \right\}, \tag{7.1}$$

if and only if $\mathcal{B} = \{\mathbf{w} \in \mathcal{L}_1^{loc}(\mathbb{R},\mathbb{R}^n) \mid \tilde{R}(\frac{d}{dt})\mathbf{w} = 0\}$ with $\tilde{R} = \begin{bmatrix} R_1 & -R_2 \end{bmatrix} T^{-1}$. We then use standard forms for polynomial matrices (see Appendix A) to construct representations of the behavior $\mathcal{B}$ in (2.1) in the form indicated in Lemma 7.1. These representations can be efficiently computed using symbolic algebra programs. By combining the representation (7.1) with the realizations of $\mathcal{B}_{i/o}$ in Theorems 4.1 and 5.3, we obtain realizations for the general behavior $\mathcal{B}$ in (2.1).

Proof (Lemma 7.1) Suppose that $\mathbf{w} \in \mathcal{B}$ where $\mathcal{B}$ satisfies (7.1). Then $\mathbf{w} = T\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix}$ for some $\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \in \mathcal{B}_{i/o}$, whence $\mathbf{w} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$. Furthermore, $\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} = T^{-1}\mathbf{w}$, and so $\mathbf{w}$ satisfies $\tilde{R}(\frac{d}{dt})\mathbf{w} = 0$. Now, let $\mathbf{w} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^n)$, satisfy $\tilde{R}(\frac{d}{dt})\mathbf{w} = 0$. Then define $\mathbf{y} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^m)$ and $\mathbf{u} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^{n-m})$ as $\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} := T^{-1}\mathbf{w}$, and we find $R_1(\frac{d}{dt})\mathbf{y} - R_2(\frac{d}{dt})\mathbf{u} = \tilde{R}(\frac{d}{dt})\mathbf{w} = 0$, so $\operatorname{col}\begin{pmatrix}\mathbf{y} & \mathbf{u}\end{pmatrix} \in \mathcal{B}_{i/o}$. ∎

We now demonstrate how to construct representations of $\mathcal{B}$ in (2.1) in the form (7.1) using either the upper echelon or row reduced form for a polynomial matrix (see Appendix A). We first note that both of these forms yield a unimodular U and $\tilde{R} \in \mathbb{R}^{m \times n}[\xi]$ as in (A.1), where m is the rank of $R(\lambda)$, and $\tilde{R}(\lambda)$ has full row-rank (equal to m), for almost all $\lambda \in \mathbb{C}$. From Section 2, we conclude that $\mathcal{B}$ satisfies:

$$(7.2)\quad \mathcal{B} = \left\{ \mathbf{w} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \;\middle|\; \begin{array}{l} \tilde{R}(\frac{d}{dt})\mathbf{w} = 0, \quad \tilde{R} \in \mathbb{R}^{m \times n}[\xi], \\ \text{and } \tilde{R}(\lambda) \text{ has full row rank for almost all } \lambda \in \mathbb{C} \end{array} \right\}.$$

It follows that there exists a non-singular matrix $T \in \mathbb{R}^{n \times n}$ with a partitioning $\begin{bmatrix}T_1 & T_2\end{bmatrix}$ such that $\tilde{R}\begin{bmatrix}T_1 & T_2\end{bmatrix} = \begin{bmatrix}R_1 & -R_2\end{bmatrix}$, where $R_1 := \tilde{R}T_1 \in \mathbb{R}^{m \times m}[\xi]$ and $R_1(\lambda)$ is non-singular for almost all $\lambda \in \mathbb{C}$. From Lemma 7.1, we conclude that $\mathcal{B}$ has a representation as in (7.1), where $\mathcal{B}_{i/o}$ is as in (2.2) with $R_1 = \tilde{R}T_1$ and $R_2 = -\tilde{R}T_2$.

The matrix T can be chosen as a permutation matrix using one of the following two methods (note that this is not an exhaustative survey of all the possibilities): (i) let $\tilde{R}$ be in upper echelon form and select the columns from $\tilde{R}$ in order to make R_1 upper triangular; (ii) let $\tilde{R}$ be in row reduced form and select the columns from $\tilde{R}$ so the leading coefficient matrix of R_1 is non-singular. In both case (i) and (ii) then, in the preceding paragraph, we let T_1 be a matrix which selects the pertinent columns in $\tilde{R}$, and T_2 a matrix which selects the remaining columns. Note that case (ii) was shown in [20, Theorem 2], and gives $R_1^{-1}R_2 \in \mathbb{R}_p^{m \times (n-m)}(\xi)$ (see also [18, Section 2]).

The methods of the previous paragraph will now be demonstrated with an example. In the behavior $\mathcal{B}$ in (2.1), we let

$$(7.3)\qquad R(\xi) = \begin{bmatrix} -4-2\xi & 2\xi^2 & 2\xi^3 & 4\xi+4 \\ -\xi^2-5\xi-5 & \xi^3+3\xi^2-1 & \xi^4+3\xi^3-\xi & 2\xi^2+8\xi+6 \\ -2\xi^2-2\xi+1 & 2\xi^3-\xi^2-\xi+1 & 2\xi^4-\xi^3-\xi^2+\xi & 4\xi^2+2\xi-2 \end{bmatrix}.$$

Using the Maple command `HermiteForm`, we obtain $U \in \mathbb{R}^{3 \times 3}[\xi]$ with $\det(U) = -1/2$, and an upper echelon matrix $\tilde{R} \in \mathbb{R}^{2 \times 4}[\xi]$ as in (A.1), where

$$(7.4)\qquad U(\xi) = \begin{bmatrix} -\frac{1}{2}\xi-\frac{3}{2} & 1 & 0 \\ -\frac{1}{2}\xi^2-\frac{5}{2}\xi-\frac{5}{2} & \xi+2 & 0 \\ \frac{1}{2}\xi^2-1 & -\xi+1 & 1 \end{bmatrix}, \text{ and } \tilde{R}(\xi) = \begin{bmatrix} 1 & -1 & -\xi & 0 \\ 0 & \xi^2-\xi-2 & \xi^3-\xi^2-2\xi & 2\xi+2 \end{bmatrix}.$$

The first two columns in $\tilde{R}$ form an upper triangular matrix, so we select these for the matrix R_1. Thus, in (7.1), we take $T = I_4$, and $R_1, R_2 \in \mathbb{R}^{2 \times 2}[\xi]$ are obtained from the partition $\tilde{R} = \begin{bmatrix}R_1 & -R_2\end{bmatrix}$.

Next, we note that the first, second, and fourth columns of the leading coefficient matrix for R are all zero, so R is not in row reduced form. Using the Maple command `RowReducedForm`, we obtain

$$(7.5)\qquad U(\xi) = \begin{bmatrix} -\frac{1}{2}\xi^3-\frac{3}{2}\xi^2+\frac{1}{2} & \xi^2 & 0 \\ -\frac{1}{2}\xi-\frac{3}{2} & 1 & 0 \\ 8\xi^2-16 & -16\xi+16 & 16 \end{bmatrix}, \text{ and } \tilde{R}(\xi) = \begin{bmatrix} \xi^2-\xi-2 & 0 & 0 & 2\xi+2 \\ 1 & -1 & -\xi & 0 \end{bmatrix}.$$

Evidently, the first and third columns in the leading coefficient matrix for $\tilde{R}$ are independent, and so we select these for the matrix R_1, and we form R_2 from the second and fourth columns of $\tilde{R}$. It may be verified that $R_1^{-1}R_2$ is proper in this case.

To conclude this section, we contrast the different realizations we obtain for the behavior $\mathcal{B}$ which correspond to different choices for T in (7.1). In particular, we identify properties which are invariant of the choice of T in the following theorem:

THEOREM 7.2. *Let $\mathcal{B}_{i/o}$ be as in (2.2), and let $\mathcal{B}$ be as in (7.1) with T non-singular. Then $\mathcal{B}$ is behaviorally controllable if and only if $\mathcal{B}_{i/o}$ is behaviorally controllable; and $\Delta(\mathcal{B}) = \Delta(\mathcal{B}_{i/o})$.*

Proof. From Lemma 7.1, $\mathcal{B} = \{\mathbf{w} \in \mathcal{L}_1^{\text{loc}}(\mathbb{R}, \mathbb{R}^n) \mid \tilde{R}(\frac{d}{dt})\mathbf{w} = 0\}$ where $\tilde{R} = \begin{bmatrix} R_1 & -R_2 \end{bmatrix} T^{-1}$. It follows that $\tilde{R}(\lambda)$ has full row rank for all $\lambda \in \mathbb{C}$ if and only if $\begin{bmatrix} R_1(\lambda) & -R_2(\lambda) \end{bmatrix}$ does. Thus, from Section 2, we conclude that $\mathcal{B}$ is behaviorally controllable if and only if $\mathcal{B}_{i/o}$ is.

It remains to show that $\Delta(\mathcal{B}) = \Delta(\mathcal{B}_{i/o})$. Since $\tilde{R} = \begin{bmatrix} R_1 & -R_2 \end{bmatrix} T^{-1}$, and $R_1(\lambda)$ in (2.2) is non-singular for almost all $\lambda \in \mathbb{C}$, then both $\begin{bmatrix} R_1(\lambda) & -R_2(\lambda) \end{bmatrix}$ and $\tilde{R}(\lambda)$ have full row rank for almost all $\lambda \in \mathbb{C}$, and hence $\Delta(\mathcal{B}) = \Delta(\tilde{R})$ and $\Delta(\mathcal{B}_{i/o}) = \Delta\left(\begin{bmatrix} R_1 & -R_2 \end{bmatrix}\right)$. Moreover, from the Binet Cauchy formula [8, p. 9], we obtain

$$\tilde{R}(p_1, \ldots, p_m) = \sum_{q_1 < \ldots < q_m \in \{1, \ldots, n\}} \begin{bmatrix} R_1 & -R_2 \end{bmatrix} (q_1, \ldots, q_m) \times T^{-1} \left(\begin{smallmatrix} q_1, & \ldots, & q_m \\ p_1, & \ldots, & p_m \end{smallmatrix} \right),$$

for $1 \leq p_1 < \ldots < p_m \leq n$. Since $T^{-1} \in \mathbb{R}^{n \times n}$, then $\Delta(\tilde{R}) \leq \Delta\left(\begin{bmatrix} R_1 & -R_2 \end{bmatrix}\right)$. Further, $\begin{bmatrix} R_1 & -R_2 \end{bmatrix} = \tilde{R}T$ with $T \in \mathbb{R}^{n \times n}$, and by again considering the Binet Cauchy formula we find that $\Delta\left(\begin{bmatrix} R_1 & -R_2 \end{bmatrix}\right) \leq \Delta(\tilde{R})$. Thus, $\Delta(\mathcal{B}) = \Delta(\tilde{R}) = \Delta\left(\begin{bmatrix} R_1 & -R_2 \end{bmatrix}\right) = \Delta(\mathcal{B}_{i/o})$. ☐

We conclude from Theorems 5.2 and 7.2 that the realizations of the behavior $\mathcal{B}$ obtained by Lemma 7.1 and Theorems 4.1 and 5.3 are all observable, and whether they are controllable or not is invariant of the specific choice of T in (7.1). The quantity $\delta(D) + d$, where d is the number of entries in the vector $\mathbf{x}$, is also invariant of the choice of T. However, properties such as input-output stability (or, more generally, the eigenvalues of A); the number of entries in $\mathbf{x}$; and the value of $\delta(D)$ can vary for different choices of T. To see this, note that the eigenvalues of A are the roots of the invariant polynomials of R_1, and the number of entries in $\mathbf{x}$ is the sum of the degrees of these invariant polynomials (see Section 4). Moreover, $\delta(D)$ is the McMillan degree of the pole of $R_1^{-1}R_2$ at infinity (see Section 5). All of these properties depend on the specific choice of R_1, which depends on the choice of T in (7.1).

Appendix A. Standard forms for polynomial matrices. Appropriate applications of polynomial division allow a given $R \in \mathbb{R}^{l \times n}[\xi]$ to be factorised into two particularly useful forms. Firstly, there exists a unimodular $U \in \mathbb{R}^{l \times l}[\xi]$ such that

$$UR = \operatorname{col}\left(\tilde{R} \quad 0_{(l-m) \times n}\right) \tag{A.1}$$

where $\tilde{R} \in \mathbb{R}^{m \times n}[\xi]$ is in upper echelon form [17, Theorem B.1.1], [8, Chapter VI]. Evidently, $\tilde{R}(\lambda)$ (and hence also $R(\lambda)$) has rank m for almost all $\lambda \in \mathbb{C}$. Secondly, there exist unimodular matrices $U \in \mathbb{R}^{l \times l}[\xi]$ and $V \in \mathbb{R}^{n \times n}[\xi]$ such that

$$URV = \operatorname{diag}\left(S \quad 0_{(l-m) \times (n-m)}\right) \tag{A.2}$$

$$\text{with} \quad S = \operatorname{diag}\left(\sigma_1 \quad \cdots \quad \sigma_m\right), \text{ for some } \sigma_1, \ldots, \sigma_m \in \mathbb{R}[\xi]. \tag{A.3}$$

Here, each polynomial in the sequence $\sigma_1, \ldots \sigma_m$ is non-zero and monic, and is divisible by the preceding term in the sequence, and m is the rank of $R(\lambda)$ (and also $S(\lambda)$) for almost all $\lambda \in \mathbb{C}$. This form is called the *Smith form*, and $\sigma_1, \ldots \sigma_m \in \mathbb{R}[\xi]$ are called the *invariant polynomials*, of R (see [17, Section 2.5.5] and [8, Chapter VI]).

A third useful form for polynomials matrices is the *row reduced form* [21, Theorem 2.5.14]. For a given $R \in \mathbb{R}^{l \times n}[\xi]$ with $R(\lambda)$ having rank m for almost all $\lambda \in \mathbb{C}$, there exists a unimodular U satisfying (A.1) for some $\tilde{R} \in \mathbb{R}^{m \times n}[\xi]$ whose *leading coefficient matrix* has full row rank. The leading coefficient matrix is formed from the coefficients of the terms of highest degree in each row of $\tilde{R}$.

REFERENCES

[1] A. Betser, N. Cohen, and E. Zeheb, *On solving the Lyapunov and Stein equations for a companion matrix*, Systems and Control Letters, 25 (1995), pp. 211–218.

[2] D. A. Bini and L. Gemignani, *Fast fraction-free triangularization of Bezoutians with applications to sub-resultant chain computation*, Linear Algebra Appl., 284 (1998), pp. 19–39.

[3] R. Bott and R. J. Duffin, *Impedance synthesis without use of transformers*, J. Appl. Phys., 20 (1949), p. 816.

[4] M. K. Çamlibel, J. C. Willems, and M. N. Belur, *On the dissipativity of uncontrollable systems*, in Proceedings of the 42nd IEEE Conference on Decision and Control, Hawaii, Dec. 2003.

[5] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, *State-space solutions to standard $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control problems*, IEEE Trans. on Automatic Control, 34 (1989), pp. 831 – 847.

[6] P. A. Fuhrmann and U. Helmke, *On the Use of Functional Models in Model Reduction. In Perspectives in Mathematical System Theory, Control, and Signal Processing - A Festschrift in Honor of Yutaka Yamamoto on the Occasion of his 60th Birthday*, Springer, 2010.

[7] P. A. Fuhrmann, P. Rapisarda, and Y. Yamamoto, *On the state of behaviors*, Linear Algebra Appl., 424 (2007), pp. 570 – 614.

[8] F. R. Gantmacher, *The Theory of Matrices*, vol. I, New York : Chelsea, 1980.

[9] B. L. Ho and R. E. Kalman, *Effective construction of linear, state-variable models from input/output functions*, Regelungstechnik, 12 (1966), pp. 545 – 548.

[10] T. H. Hughes, *Why RLC realizations of certain impedances require a majority of uncontrollable or unobservable modes*, Submitted to IEEE Transactions on Automatic Control, (2015).

[11] T. H. Hughes and M. C. Smith, *On the minimality and uniqueness of the Bott-Duffin realization procedure*, IEEE Trans. on Automatic Control, 59 (2014), pp. 1858–1873.

[12] ———, *Controllability of linear passive network behaviors*, Accepted in Systems and Control Letters, (2015).

[13] R. E. Kalman, *Mathematical description of linear dynamical systems*, SIAM Journal on Control, 2 (1963), pp. 152 – 192.

[14] D. W. Kammler, *Numerical evaluation of exp(tA) when A is a companion matrix*, SIAM Journal on Numerical Analysis, 15 (1978), pp. 1077–1102.

[15] I. Kaufman and P. H. Roe, *On systems described by the companion matrix*, IEEE Trans. on Automatic Control, 15 (1970), pp. 692–693.

[16] J. W. Polderman, *Proper elimination of latent variables*, Systems and Control Letters, 32 (1997), pp. 262–269.

[17] J. W. Polderman and J. C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach*, New York : Springer-Verlag, 1998.

[18] P. Rapisarda and J. C. Willems, *State maps for linear systems*, SIAM Journal on Control Optim., 35 (1997), pp. 1053 – 1091.

[19] A. J. van der Schaft and P. Rapisarda, *State maps from integration by parts*, SIAM Journal on Control Optim., 49 (2011), pp. 2415 – 2439.

[20] J. C. Willems, *From time series to linear system—Part I. finite dimensional linear time invariant systems*, Automatica, 22 (1986), pp. 561–580.

[21] W. A. Wolovich, *Linear Multivariable Systems*, New York : Springer-Verlag, 1974.

[22] D. C. Youla and P. Tissi, *N-port synthesis via reactance extraction, Part I*, IEEE International Convention Record, 14 (1966), pp. 183–205.