An Algebraic Approach to the Control of Decentralized Systems

Laurent Lessard   Sanjay Lall

Abstract

Optimal decentralized controller design is notoriously difficult, but recent research has identified large subclasses of such problems that may be convexified and thus are amenable to solution via efficient numerical methods. One recently discovered sufficient condition for convexity is *quadratic invariance* (QI). Despite the simple algebraic characterization of QI, which relates the plant and controller maps, proving convexity of the set of achievable closed-loop maps requires tools from functional analysis. In this work, we present a new formulation of quadratic invariance that is purely algebraic. While our results are similar in flavor to those from traditional QI theory, they do not follow from that body of work. Furthermore, they are applicable to new types of systems that are difficult to treat using functional analysis. Examples discussed include rational transfer matrices, systems with delays, and multidimensional systems.

I Introduction

The problem of designing control systems where multiple controllers are interconnected over a network to control a collection of interconnected plants is long-standing and difficult [4,31]. *Quadratic invariance* is a mathematical condition which, when it holds, allows one to bring to bear the tools of Youla parameterization to find optimal controllers [23,24]. A network system has the requisite quadratic invariance under a surprisingly wide set of circumstances. These include cases where the controllers can communicate more quickly than the plant dynamics propagate through the network [22].

There is a large and diverse body of literature addressing decentralized control theory and specifically conditions that make a problem more tractable in some sense. The seminal work of Ho and Chu [9] develops the *partial nestedness* condition under which there exists an optimal decentralized controller that is linear. More recently, Qi, Salapaka, et al. identified many different decentralized control architectures that may be cast as tractable optimization problems [18]. LMI formulations of distributed control problems are developed in [6, 14]. Stabilization was fully characterized for all QI problems in [25]. Explicit state-space solutions were also found for classes of delayed problems [12], poset-causal problems [26], and two-player output-feedback problems [17].

There have been relatively few works treating decentralized control from a purely algebraic perspective. One recent example is the work of Shin and Lall [27], where elimination theory is used to express solutions to decentralized control problems as projections of semialgebraic sets. Quadratic invariance was first treated using an analytic framework in [23,24]. The aim of this paper is to address an algebraic treatment of quadratic invariance. The consequence of this is not only a new proof of existing results in some cases but also an extension of these results to a significantly different class of models. Instead of requiring analytic properties of our system model, we will require algebraic ones. For example, [24] requires that the set of allowable controllers be a closed inert subspace,

whereas in this work we require that it be a module. The class of systems covered in this paper includes multidimensional systems, which are not covered in existing works. This is discussed in Section VI-C.

Many topics in control have historically been treated from both analytic and algebraic viewpoints. As early as 1965, Kalman proposed the use of modules as the natural framework in which to represent linear state-space systems [10]. When systems are viewed as maps on signal spaces, one has many choices. If one represents systems as transfer functions, then one can either consider the generality of transfer functions in a Hardy space and use analytic methods to prove results, or one can consider formal power series or rational functions and use algebraic methods. Often, the two frameworks use very different proof techniques, which provide different insights and ranges of applicability. This is a fundamental choice in how we represent the basic objects [11]. This dichotomy exists in many facets of the control systems literature. For example, spectral factorization is easily considered from an algebraic perspective. The Riesz-Fejér theorem states that a trigonometric polynomial which is nonnegative on the circle may be factored into the product of two polynomials, one of which is holomorphic inside the disc and the other outside [19]. This is the fundamental algebraic version of the discrete-time SISO spectral factorization result. For comparison, the analytic version of this result is commonly known as Wiener-Hopf factorization [30]. Of course, the same choice of frameworks exists beyond factorization. The theory of stabilization as introduced by Youla [32] was developed both algebraically [29] and analytically [28]. The idea of algebraic representations have also proven useful in areas such as realization theory [2], model reduction [3], and nonlinear systems theory [7].

The work in this paper is based on preliminary results that first appeared in [15, 16]. Unlike these early works, all invariance results in the present work include both necessary and sufficient conditions, and all the proofs are purely algebraic. The paper is organized as follows. The remainder of the introduction gives an overview of quadratic invariance and existing analytic results. Invariance results are proven and discussed for matrices, rings and fields, and rational functions in Sections II, IV, and V, respectively. We present some illustrative examples in Section VI and we summarize our contributions in Section VIII.

I-A   Quadratic invariance

We have adopted the notation convention from [23, 24] to make the works readily comparable. Given a plant G, which is a map from an input space $\mathcal{U}$ to an output space $\mathcal{Y}$, we seek to design a controller $K : \mathcal{Y} \to \mathcal{U}$ that achieves desirable performance when connected in feedback with G. The main object of interest is the function $h : M \to M$ given by

$$h(K) = -K(I - GK)^{-1}$$

Here, the domain M is the set of maps $K : \mathcal{Y} \to \mathcal{U}$ such that $I - GK$ is invertible. The image of h is again M because h is an *involution*. That is, $h(h(K)) = K$. We will be more specific shortly about the nature of the spaces $\mathcal{U}$ and $\mathcal{Y}$ (and consequently the maps G and K).

The motivation for studying h is that it is a linear fractional transform that occurs in feedback control. Consider for example the four-block plant of Figure 1.
For simplicity, assume for now that $S \subseteq M$. In Figure 1, the set of achievable closed-loop maps $w \mapsto z$ subject to K belonging to some set S is given by

$$\mathcal{C} = \{P_{11} - P_{12}h(K)P_{21} \mid K \in S\}$$

Selecting a controller that optimizes some closed-loop performance metric is equivalent to selecting the best $T \in \mathcal{C}$ and then finding the $K \in S$ that yields this T.
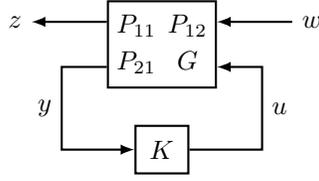
Figure 1: Four-block plant with controller in feedback

Roughly, the works [23, 24] give a necessary and sufficient condition such that $h(S) = S$. If this condition holds, then $\mathcal{C} = \{P_{11} - P_{12}KP_{21} \mid K \in S\}$, and so the set of achievable closed-loop maps is affine and easily searchable. The condition is called **quadratic invariance**, and a generic definition is given below.

Definition 1 (Quadratic invariance). *We say that the set S is quadratically invariant (QI) under G if for all $K \in S$, we have $KGK \in S$.*

In [23], the input and output spaces are Banach spaces, and G and K are bounded linear operators. In [24], the extended spaces ℓ_{2e} and L_{2e} are used and the associated maps are then continuous linear operators. In this work we use different spaces still, but the generic definition of quadratic invariance remains the same.

We now state the main results from [23, 24]. Additional notation and terminology is defined after each theorem statement.

Theorem 2 (see [23]). *Suppose $\mathcal{Y}$ and $\mathcal{U}$ are Banach spaces, $G \in L(\mathcal{U}, \mathcal{Y})$ and $S \subseteq L(\mathcal{Y}, \mathcal{U})$ is a closed subspace. Further suppose that $N \cap S = M \cap S$. Then*

$$S \text{ is QI with respect to } G \quad \Longleftrightarrow \quad h(S \cap M) = S \cap M$$

In Theorem 2, $L(\cdot, \cdot)$ denotes the set of bounded linear operator from the first argument to the second. Also, N is the set of $K \in L(\mathcal{Y}, \mathcal{U})$ such that $1 \in \rho_{\mathrm{uc}}(GK)$, the unbounded connected component of the resolvent set of GK. The condition that $N \cap S = M \cap S$ is admittedly technical in nature, but the result of the theorem is very simple; quadratic invariance is equivalent to S being invariant under h.

Theorem 3 (see [24]). *Suppose $G \in \mathcal{L}(L_{2e}^n, L_{2e}^m)$ and $S \subseteq \mathcal{L}(L_{2e}^m, L_{2e}^n)$ is an inert closed subspace. Then*

$$S \text{ is QI with respect to } G \quad \Longleftrightarrow \quad h(S) = S$$

In Theorem 3, $\mathcal{L}(\cdot, \cdot)$ denotes the set of continuous linear maps from the first argument to the second. The requirement that S be *inert* means that the impulse response matrix of GK must be entry-wise bounded over every finite time interval for all $K \in S$. Among other things, this technical condition guarantees that $I - GK$ is always invertible, and so h is well-defined over all of S. An analogous result to Theorem 3 in which L_{2e} is replaced by ℓ_{2e} is also provided in [24].

It is interesting to note that both sides of the equivalences proved in Theorems 2 and 3 are purely algebraic statements. In other words, they can be stated in terms of a finite number of algebraic operations (addition, multiplication, inversion). This seems at odds with the technical assumptions required in the theorems. For example, S being a closed subspace means that S should contain all of its limit points. This is an analytic concept requiring an underlying norm or a topology at the very least.

This observation is the starting point for this work, where we show that *invariance results* akin to Theorems 2 and 3 can be obtained in a purely algebraic setting without

requiring anything more than well-defined addition, multiplication, and inversion. This makes rings and fields the natural objects to work with, and we will discuss them at greater length in Section III. In Section VI we give three specific settings where these algebraic tools offer a natural framework for modeling control systems. These are the cases of sparse controllers, networks with delays, and multidimensional systems.

II   The matrix case

The real matrix case is an example that illustrates when quadratic invariance may be treated either analytically or algebraically. In this section, we present the invariance result in the matrix case, and give an outline of the proof using both the existing analytic approach [23, 24] and the algebraic approach that is expanded upon in more detail in later sections of this work. We present the proofs in sufficient detail to highlight the mathematical machinery being used, but we skip over less relevant details in the interest of clarity.

Theorem 4 (QI for matrices). *Suppose $G \in \mathbb{R}^{m \times n}$ and $S \subseteq \mathbb{R}^{n \times m}$ is a subspace. Then the following holds.*

$$S \text{ is QI with respect to } G \quad \Longleftrightarrow \quad h(S \cap M) = S \cap M$$

Proof. We outline a proof of the forward direction ($\Longrightarrow$). If S is QI with respect to G, then by definition we have $KGK \in S$ for all $K \in S$. The first step is to show that $K(GK)^i \in S$ for $i = 1, 2, \ldots$ as well. This can be proven by induction using the identity

$$K(GK)^{i+1} = \frac{1}{2}\Big(\big(K + K(GK)^i\big)G\big(K + K(GK)^i\big) - KGK - \big(K(GK)^i\big)G\big(K(GK)^i\big)\Big)$$

Next, we examine the function $h(K) = K(I - GK)^{-1}$ when $K \in S \cap M$. It suffices to show that $h(K) \in S$, since h is involutive. We prove this result first via an analytic approach similar to the one used in [23], and then using an algebraic approach.

The *analytic approach* is to use an infinite series expansion. For $\alpha \in \mathbb{C}$, we have the following convergent series.

$$K(I - \alpha GK)^{-1} = \sum_{i=0}^{\infty} K(GK)^i \, \alpha^i \quad \text{for } |\alpha| < \frac{1}{\|GK\|} \tag{1}$$

Since $K(GK)^i \in S$ for all $K \in S$, and S is a finite-dimensional subspace and therefore closed, the infinite sum converges to an element of S. Using an analytic continuation argument [23], one can show that $K(I - \alpha GK)^{-1} \in S$ for all α such that $\det(I - \alpha GK) \neq 0$. It then follows that $h(K) \in S$ for all $K \in S \cap M$, as required.

The *algebraic approach* is to use a finite series expansion. Pick some $K \in S$ such that $(I - GK) \in \mathbb{R}^{m \times m}$ is invertible. By the Cayley-Hamilton theorem, there exist $p_0, \ldots, p_{m-1} \in \mathbb{R}$ such that

$$(I - GK)^{-1} = p_0 I + p_1(I - GK) + \cdots + p_{m-1}(I - GK)^{m-1}$$

Expanding and collecting like powers of GK, we find that

$$K(I - GK)^{-1} = q_0 K + q_1 KGK + \cdots + q_{m-1}K(GK)^{m-1}$$

for some $q_0, \ldots, q_{m-1} \in \mathbb{R}$. Once again, $K(GK)^i \in S$ for all $K \in S$, so every term in this finite sum belongs to the subspace S and therefore $K(I - GK)^{-1} \in S$. The difference is

4

that we did not require an analytic continuation, nor did we make use of the fact that S is closed.

The converse direction ($\Longleftarrow$) can also be proven either using an analytic argument as in [23], or using an algebraic argument as we will develop in Section IV. ∎

Notice that we give two proofs of the forward direction of Theorem 4, one analytic and the other algebraic. The analytic approach is based on convergence and hence depends on the topology. Since this particular result is only stated for matrices, the choice of topology does not matter, but for more general convolution operators on infinite signal spaces the topology has a significant effect on the applicability of the result and the technical machinery required to effect the proof. However, the algebraic approach is much more simple, and only relies on addition, multiplication and inversion.

The main results of this paper, given in Sections IV and V, generalize and expand upon the algebraic approach used above to matrices with entries that belong to a commutative ring R.

III   Algebraic preliminaries

The fundamental algebraic properties we wish to capture are simply addition and multiplication. This leads naturally to rings and fields, which are a fundamental building block of abstract algebra. These concepts also provide the framework which is commonly used to state many widely-used results in control theory. For example, the set of real-rational transfer functions is a field, and the subset of proper ones is a ring. This viewpoint can be extremely useful, for example, when parameterizing all stabilizing controllers [29]. We refer the reader to [13] for an introduction to these concepts.

We now explain some of the conventions used throughout this paper. The integers, reals, rationals, and complex numbers are denoted by $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$, and $\mathbb{C}$, respectively. We use R to denote an arbitrary commutative ring with identity, and $\mathbb{F}$ to denote an arbitrary field. The additive and multiplicative identity elements of R are denoted 0_R and 1_R, respectively, but we will often omit the subscript when it is clear by context. An invertible element of R is called a **unit**, and the set of units is written as $U(R)$. We write $R[x]$ and $\mathbb{F}(x)$ to respectively denote the ring of polynomials and the field of rational functions in the indeterminate x. If $H \subseteq R$ is an ideal, we write $H \lhd R$. If $M \lhd R$ is maximal ideal, the associated quotient ring R/M is a field, and is called the **residue field**. Finally, we make use of the notion of an R-module, which is the generalization of a vector space when the scalars belong to a ring R rather than a field $\mathbb{F}$.

In this paper, we consider finite matrices with elements that belong to R. Much of the familiar linear algebra theory carries over to this more general setting. We refer the reader to [5] for introduction to abstract linear algebra. We write $R^{m \times n}$ to mean the set of $m \times n$ matrices with entries in R. Matrix multiplication between matrices of compatible dimensions is defined in the standard way. When the matrices are square, we write $M_n(R) = R^{n \times n}$, which is a ring. The identity matrix is denoted I_R, that is, the matrix whose diagonal and off-diagonal entries are 1_R and 0_R, respectively.

Many concepts from matrix theory carry over to the more general setting. Specifically, if $A \in M_n(R)$, the determinant $\det : M_n(R) \to R$ is defined by the conventional Laplace expansion. The adjugate (or classical adjoint) $\mathrm{adj} : M_n(R) \to M_n(R)$ also makes sense, as it is defined in terms of determinants of submatrices. The fundamental identity for adjugates holds as well, namely

$$A \, \mathrm{adj}(A) = \mathrm{adj}(A)A = \det(A)I_R. \tag{2}$$

The matrix $A \in M_n(R)$ is invertible if and only if $\det(A) \in U(R)$. In this case, the inverse is unique, and is given by
$$A^{-1} = (\det(A))^{-1} \operatorname{adj}(A)$$

For an introduction to the adjugate and associated results, we refer the reader to [20]. We now state a fundamental result.

Proposition 5. *Suppose $A \in M_n(R)$. Let $p_A \in R[x]$ be the characteristic polynomial of A, given by $p_A(x) = \det(A - xI)$. Suppose p_A has the form*

$$p_A(x) = p_0 + p_1 x + \cdots + p_n x^n$$

Then the following equations hold in $M_n(R)$.

(i) $p_0 I + p_1 A + \cdots + p_n A^n = 0$

(ii) $\operatorname{adj}(A) = -(p_1 I + p_2 A + \cdots + p_n A^{n-1})$

Proof. See Section VII. ∎

Item (i) in Proposition 5 is commonly known as the Cayley-Hamilton theorem. This result plays an important role in our approach because it enables us to express quantities such as $(I - GK)^{-1}$ as finite sums.

IV   Invariance for rings and fields

In this section, we take the notion of quadratic invariance discussed in the introduction and show how it fits into the framework of matrices over commutative rings or fields. These results generalize the algebraic invariance result for real matrices from Section II. Complete proofs for all the results of this section are given in Section VII.

Our first main invariance result holds over matrices whose entries belong to an arbitrary commutative ring R with identity. Terminology is explained after the theorem statement.

Theorem 6 (QI for rings). *Suppose $G \in R^{m \times n}$ and $S \subseteq R^{n \times m}$ is an R-module.*

1) *If $2_R \in U(R)$, then*

$$S \text{ is QI with respect to } G \quad \Longrightarrow \quad K \operatorname{adj}(I_R - GK) \in S \text{ for all } K \in S$$

2) *If every residue field of R has at least $\min(m, n) + 1$ elements, then*

$$S \text{ is QI with respect to } G \quad \Longleftarrow \quad K \operatorname{adj}(I_R - GK) \in S \text{ for all } K \in S$$

The notion of R-module is analogous to that of a subspace. That is, S contains all linear combinations of its elements, where the linear combinations have coefficients in R.

Theorem 6 contains several technical conditions which we will now explain. For the ($\Longrightarrow$) result, the condition $2_R \in U(R)$ means that $2_R := 1_R + 1_R$ must be a unit. We will now show that this condition is necessary by providing a counterexample. The condition is satisfied for example in the ring of rationals $\mathbb{Q}$, but not in the ring of integers $\mathbb{Z}$. Consider therefore the following integer example.

$$S = \left\{ \begin{bmatrix} 2x & y & z \\ y & z & 0 \\ z & 0 & 0 \end{bmatrix} \middle| \ x, y, z \in \mathbb{Z} \right\}, \quad G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

It is straightforward to check that S is a $\mathbb{Z}$-module, and is quadratically invariant with respect to G. Now consider the following particular element of S.

$$K_0 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \in S$$

and so

$$K_0 \operatorname{adj}(I - GK_0) = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \notin S$$

Therefore the first part of Theorem 6 does not hold. For more details on why this is so, refer to the proof of Theorem 6 in Section VII.

For the ($\Longleftarrow$) result, a residue field is the field obtained by taking the quotient R/M for some maximal ideal $M \lhd R$. For example, in the ring of integers $\mathbb{Z}$, the maximal ideals are the sets $p\mathbb{Z}$ for some prime p. So for $p = 2$, the associated ideal is the set of even integers, and the residue field is $\mathbb{Z}/2\mathbb{Z}$; the integers modulo 2. This field only has two elements and so the conditions of the theorem would not be satisfied for matrices with at least two rows or two columns.

We now specialize the above invariance result to fields. Axiomatically, a field is simply a ring for which every nonzero element is a unit. The results of this section hold for an arbitrary field $\mathbb{F}$. We begin by stating the invariance result, and then we explain the differences between the field and ring cases. In particular, several concepts become simpler when the ring in question is a field.

Theorem 7 (QI for fields). *Suppose $G \in \mathbb{F}^{m \times n}$ and $S \subseteq \mathbb{F}^{n \times m}$ is a subspace over $\mathbb{F}$. Further suppose that $\mathbb{F}$ contains at least $2\min(m,n)+1$ distinct elements and $\operatorname{char} \mathbb{F} \neq 2$.*

$$S \text{ is QI with respect to } G \quad \Longleftrightarrow \quad h(S \cap M) = S \cap M$$

where $M = \{K \in \mathbb{F}^{n \times m} \mid \det(I - GK) \neq 0\}$.

The *characteristic* of a field $\mathbb{F}$, denoted $\operatorname{char} \mathbb{F}$ is the smallest k such that $\overbrace{1 + \cdots + 1}^{k \text{ times}} = 0$. When there is no such k, then we say $\operatorname{char} \mathbb{F} = 0$. Note that requiring $\operatorname{char} \mathbb{F} \neq 2$ is the same as the condition that $2_R \in U(R)$ when the ring R is specialized to the field $\mathbb{F}$.

In the real-number case $\mathbb{F} = \mathbb{R}$, both technical assumptions are always satisfied because $\mathbb{R}$ has infinitely many elements and $\operatorname{char} \mathbb{R} = 0$. We then precisely recover Theorem 4. Note that Theorem 7 may also be applied to finite fields such as $\mathbb{Z}/p\mathbb{Z}$, the field of integers modulo p.

V   Invariance for rationals

In this section, we specialize the ring and field invariance results of Section IV to rational functions in multiple variables. This leads to quadratic invariance results without any technical requirement on S such as closure or the existence of limits. As we shall see in Section VI, this framework can accommodate systems with delays or spatiotemporal systems.

Let $\mathbb{R}(x)$ be the set of rational functions in the indeterminate x with coefficients in $\mathbb{R}$. Because $\mathbb{R}(x)$ is a field, we may apply Theorem 7. We obtain the following result.

Theorem 8 (QI for rationals). *Suppose $G \in \mathbb{R}(x)^{m \times n}$, and $S \subseteq \mathbb{R}(x)^{n \times m}$ is an $\mathbb{R}(x)$-module.*

$$S \text{ is QI with respect to } G \quad \Longleftrightarrow \quad h(S \cap M) = S \cap M$$

Theorem 8 is the simplest algebraic result for quadratic invariance of rational functions, and provides the technical basis for the remainder of this paper. However, it is not directly applicable to most control systems, because for physical models one typically has the constraint that the system G is proper. This applies for example if G is a transfer function that represents a causal time-invariant system, and we seek a controller that is also causal and time-invariant.

Let $\mathbb{R}(s)_{\mathrm{p}}$ be the set of proper rational functions in the indeterminate s, and let $\mathbb{R}(s)_{\mathrm{sp}} \subset \mathbb{R}(s)_{\mathrm{p}}$ denote the strictly proper rationals. Note that proper rationals are a ring rather than a field, because the inverse of a proper rational function is generally not proper. The result is given below.

Theorem 9 (QI for proper rationals). *Suppose $G \in \mathbb{R}(s)_{\mathrm{sp}}^{m \times n}$, and $S \subseteq \mathbb{R}(s)_{\mathrm{p}}^{n \times m}$ is an $\mathbb{R}(s)_{\mathrm{p}}$-module.*

$$S \text{ is QI with respect to } G \quad \Longleftrightarrow \quad h(S) = S$$

We now extend the rational results of Theorems 8 and 9 to rational functions of multiple variables with mixed properness constraints. This will allow this approach to be applied to two additional classes of systems. The first is networks of linear systems interconnected by delays, and the second is multidimensional systems. These are discussed in Section VI. Specifically, define the sets of indeterminates $\mathbf{x} = (x_1, \ldots, x_\ell)$ and $\mathbf{s} = (s_1, \ldots, s_k)$. We are interested in the ring of multivariate rationals $\mathbb{R}(\mathbf{x}, \mathbf{s})$ where we have imposed a properness constraint on each of the $s_i \in \mathbf{s}$. Note that this is not the same as the rational function itself being proper. For example,

$$g = \frac{s_1 s_2 s_3}{s_1^2 + 2s_2 + s_3}$$

is proper in each of the variables s_1, s_2, s_3, but is not proper by the standard definition since the degree of the numerator is larger than that of the denominator. We will use the subscripts p or sp to apply individually to each of the $s_i \in \mathbf{s}$ while ignoring the $x_i \in \mathbf{x}$. Therefore, $g \in \mathbb{R}(\mathbf{x}, \mathbf{s})_{\mathrm{p}}$. The multivariate rational invariance result is given below.

Theorem 10 (QI for multivariate rationals). *Suppose that $G \in \mathbb{R}(\mathbf{x}, \mathbf{s})_{\mathrm{sp}}^{m \times n}$, and $S \subseteq \mathbb{R}(\mathbf{x}, \mathbf{s})_{\mathrm{p}}^{n \times m}$ is an $\mathbb{R}(\mathbf{x}, \mathbf{s})_{\mathrm{p}}$-module.*

$$S \text{ is QI with respect to } G \quad \Longleftrightarrow \quad h(S) = S$$

In Section VI, we give an example of a class of systems that can be represented by a multivariate rational function with mixed properness constraints such as those used in Theorem 10.

The invariance results of Section IV only rely on the algebraic properties of the objects involved, so our results may be applied to a variety of examples beyond the ones mentioned in this section. As a simple example, we may replace $\mathbb{R}$ by $\mathbb{Q}$ or $\mathbb{C}$ in Theorems 8–10.

VI   Examples

In this section, we show some examples of problems that can be modeled using our algebraic framework. The purpose is to illustrate that the constraint that S be an R_{p}-module occurs frequently and in a variety of different situations.

VI-A   Sparse controllers

The simplest class of systems that we can analyze are systems with rational transfer functions subject to controllers with sparsity constraints. If every nonzero entry in the controller is required to be a proper rational function in $\mathbb{R}(s)_{\mathrm{p}}$, it is clear that the set S of admissible controllers is an $\mathbb{R}(s)_{\mathrm{p}}$-module.

VI-B   Network with delays

Consider a distributed system where the subsystems affect one another via delay constraints. We wish to design a decentralized controller subject to communication delay constraints between subcontrollers.

Consider the simple example of two plants, each with their own controller. We represent the plants and their associated controllers by the transfer functions $(G_i(s), K_i(s))$. Suppose the controllers communicate with each other using a bilateral network that taxes all transmissions with a delay $d = e^{-s\tau}$. The example is illustrated in Figure 2 For sim-
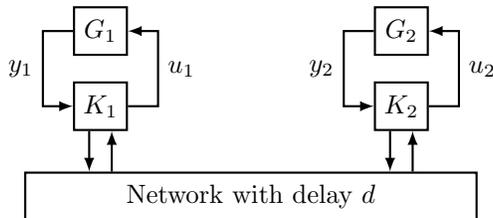

Figure 2: Simple example consisting of two plants with networked controllers.

plicity, suppose each controller transmits all the measurements it receives to the other controller. Then the global plant and controller are characterized by the maps

$$
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12}d \\ K_{21}d & K_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}
$$

Such an architecture is QI, and more detailed topologies of this type were studied in [22, 24]. In general, the algebraic framework allows us to treat scenarios where the plant G and controller K are rational functions in s and d. The constraint $K \in \mathbb{R}(s, d)_{\mathrm{p}}$, where properness is enforced on s and d independently, naturally guarantees that negative delays are forbidden, thus enforcing causality.

Define the *delay* of a transfer function as the difference between the degree of d in its denominator and numerator. For example,

$$
\text{delay}\left(\frac{1}{sd + 2}\right) = 1 \quad \text{and} \quad \text{delay}\left(\frac{s + d^2}{s^2 d + d^5}\right) = 3
$$

As a convention, $\text{delay}(0) = \infty$. We can impose delay constraints on the controller using a set of the form

$$
S = \{K \in \mathbb{R}(s, d)_{\mathrm{p}} \mid \text{delay}(K_{ij}) \geq a_{ij}\}
$$

where $a_{ij} \geq 0$ is the minimum delay (in multiples of d) between subcontrollers i and j. One can verify that S is an $\mathbb{R}(s, d)_{\mathrm{p}}$-module, and so we may apply Theorem 10 to deduce an invariance condition. Similar results proved using very different methods can be found in [22].

VI-C   Multidimensional systems

In many cases, we are interested in modeling and control of systems whose states, inputs or outputs may be functions of a spatial independent variable, such as in control of continuum mechanical models such as fluids or elastic solids. While the dynamics may be readily described by linear operator models, analysis of such models has typically been performed using the tools of C_0 semigroup theory [1]. Taking Fourier transforms spatially and Laplace transforms temporally, one arrives at an algebraic formulation of

such systems [2, 21], where a linear system is described by a rational function of two or more independent frequency variables. The properness requirement that arises from causality is then only required with respect to the temporal frequency variable. This notion of multivariate transfer functions is used to represent spatiotemporal dynamics in a variety of important papers [3,6,8,21]. Using this framework, we are able to circumvent the analytical requirements and still explicitly construct the set of closed-loop maps for such systems. This class of systems is not addressed by existing results on quadratic invariance, and in particular it is not covered by the results of [24]. Our approach allows one to simply describe the set of achievable closed-loop maps for such systems in both the centralized and decentralized cases. In general, our framework allows us to consider transfer functions in two sets of variables $R = \mathbb{R}(x_1, \ldots, x_\ell, s_1, \ldots, s_k)$ where we impose properness on the s_i but not on the x_i.

It is however worth noting that synthesis of the optimal controller for such multidimensional systems is a challenging problem, even in the centralized case. Therefore one cannot simply combine the results in this paper with existing exact synthesis formulae or tools from the centralized case, and synthesize decentralized multidimensional controllers.

VII   Proofs of main results

Proof of Proposition 5. Apply the identity (2) to $A - xI$, which we view as an element of $M_n(R[x])$, and obtain

$$(A - xI) \operatorname{adj}(A - xI) = (p_0 + p_1 x + \cdots + p_n x^n)I \tag{3}$$

The adjugate is defined in terms of minors, so each entry of $\operatorname{adj}(A - xI)$ is an element of $R[x]$ of degree at most $n - 1$. Because of the ring isomorphism $M_n(R[x]) \cong (M_n(R))[x]$ (see [5, §III.C] for a proof), we may write $\operatorname{adj}(A - xI) = B_0 + B_1 x + \cdots + B_{n-1} x^{n-1}$ for some $B_i \in M_n(R)$. Substituting into (3) and viewing the result as as a polynomial identity in $(M_n(R))[x]$, we obtain

$$(A - Ix)(B_0 + B_1 x + \cdots + B_{n-1} x^{n-1}) = p_0 I + p_1 I x + \cdots + p_n I x^n \tag{4}$$

Expanding (4) and comparing coefficients, we obtain

$$AB_0 = p_0 I$$
$$AB_k - B_{k-1} = p_k I \qquad \text{for } k = 1, \ldots, n-1$$
$$-B_{n-1} = p_n I$$

Left-multiplying each p_k equation by A^k and summing all equations, we obtain (i). Similarly, left-multiplying each p_k equation by A^{k-1} for $k \geq 1$ and summing, we obtain

$$p_1 + p_2 A + \cdots + p_n A^{n-1} = -B_0 = -\operatorname{adj}(A)$$

which is the statement of (ii).   ∎

VII-A   Invariance for rings

In the following subsection, R is an arbitrary commutative ring with identity. We begin with a lemma that allows us conclude that if $K \in S$ and $KGK \in S$, then we have $K(GK)^i \in S$ for $i = 1, 2, 3, \ldots$.

Lemma 11. *Suppose $G \in R^{m \times n}$ and $S \subseteq R^{n \times m}$ is an R-module. Further suppose that $2_R \in U(R)$. If S is quadratically invariant with respect to G, then for all $K \in S$:*

$$K(GK)^i \in S \qquad for \; i = 1, 2, \ldots$$

Proof. The result follows by induction, using the identity:

$$K(GK)^{i+1} = 2_R^{-1}\Big(\big(K + K(GK)^i\big)G\big(K + K(GK)^i\big) - KGK - \big(K(GK)^i\big)G\big(K(GK)^i\big)\Big)$$

where 2_R^{-1} is the multiplicative inverse of $2_R = 1_R + 1_R$, which exists by assumption. ■

Note that Lemma 11 requires that $2_R \in U(R)$. As shown in the first example of Section IV, this requirement is necessary. If we strengthen the notion of quadratic invariance to instead require that $K_1 G K_2 \in S$ for all $K_1, K_2 \in S$, then the conclusion of Lemma 11 is trivial and the assumption $2_R \in U(R)$ is no longer required.

We will require an important property of polynomials. Specifically, we will need conditions under which a polynomial is uniquely specified by the values it takes on at finitely many points. For a real polynomial $f \in \mathbb{R}[x]$, the result is well-known. If we have $f(x) = 0$ for all $x \in \mathbb{R}$, then $f = 0$. In other words, all coefficients of f are zeros and thus f is the zero polynomial. The same result does not hold when we replace $\mathbb{R}$ by a general field $\mathbb{F}$ or by a commutative ring R. As a simple example, consider $f = x + x^2 \in (\mathbb{Z}/2\mathbb{Z})[x]$, a polynomial with coefficients in the integers modulo 2. Then clearly $f(0) = f(1) = 0$, yet $f \neq 0$.

Polynomials are closely related to *Vandermonde matrices*, which we now define. The $N \times n$ Vandermonde matrix generated by $r_1, \ldots, r_N \in R$ is defined as

$$V = \begin{bmatrix} 1 & r_1 & \cdots & r_1^{n-1} \\ 1 & r_2 & \cdots & r_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r_N & \cdots & r_N^{n-1} \end{bmatrix} \in R^{N \times n} \tag{5}$$

So if $f(r) = a_0 + a_1 r + \cdots + a_{n-1} r^{n-1}$, then the Vandermonde matrix (5) relates the values of f evaluated at $r_1, \ldots, r_N$ to the polynomial coefficients via

$$\begin{bmatrix} f(r_1) \\ \vdots \\ f(r_N) \end{bmatrix} = V \begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix}$$

If there exists a left-invertible Vandermonde matrix V, then the coefficients a_i are uniquely determined by the values $f(r_j)$. If R is a field, we may assume without loss of generality that the Vandermonde matrix is square. The answer is given by the following proposition.

Proposition 12. *Suppose $\mathbb{F}$ is a field. The following statements are equivalent*

(i) The field $\mathbb{F}$ contains at least n distinct elements.

(ii) There exists an invertible $n \times n$ Vandermonde matrix.

(iii) There exists a left-invertible $N \times n$ Vandermonde matrix for some $N \geq n$.

The proof follows from the following well-known formula for the determinant of a square Vandermonde matrix.

$$\det V = \prod_{1 \leq i < j \leq n} (r_i - r_j)$$

The result is more complicated if R is a commutative ring with identity. In reference to Proposition 12, there are cases where (iii)$\not\Rightarrow$(ii). As an example, consider the ring $\mathbb{Z}[\beta]$ where $\beta = \frac{1}{2}(1 + \sqrt{-11})$. It is easy to check that the only units of $\mathbb{Z}[\beta]$ are ± 1. There are no 3×3 invertible Vandermonde matrices in this ring. To see why, note that if V is generated by x, y, z then $\det V = -(x - y)(y - z)(z - x)$, which is a unit if and only if each factor is a unit. Adding the factors together yields $\pm 1 \pm 1 \pm 1 = 0$, a contradiction. However, there exists a 4×3 left-invertible Vandermonde matrix in this ring. For example,

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & \beta \\ 0 & 1 & 4 & \beta^2 \end{bmatrix} \begin{bmatrix} 1 & 1 - 10\beta & 1 + \beta \\ 0 & 17 + 15\beta & -4 - \beta \\ 0 & -11 - 3\beta & 2 \\ 0 & -7 - 2\beta & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The following lemma gives a complete characterization of left-invertibility for Vandermonde matrices in a general commutative ring with identity.

Lemma 13. *Suppose R is a commutative ring with identity. The following statements are equivalent.*

(i) *Every residue field of R has at least n elements.*

(ii) *The ideal generated by the determinants of all $n \times n$ Vandermonde matrices is equal to R, the unit ideal.*

(iii) *For some $N \geq n$, there exists a left-invertible $N \times n$ Vandermonde matrix.*

Proof. We begin by showing (i) $\Longleftrightarrow$ (ii). If (i) is false, there exists some maximal ideal $M \lhd R$ such that the quotient ring R/M contains fewer than n elements. Therefore, given any set of elements $r_1, \ldots, r_n \in R$, we must have $r_i - r_j \equiv 0 \pmod{M}$ for some i, j. It follows that if V is the Vandermonde matrix generated by $r_1, \ldots, r_n$ then it satisfies

$$\det V = \prod_{1 \leq i < j \leq n} (r_i - r_j) \equiv 0 \pmod{M}$$

Therefore $\det V \in M$ for every Vandermonde matrix. So the ideal generated by all $n \times n$ Vandermonde determinants is contained in M, a proper ideal, and so cannot be the unit ideal. This shows that (ii) is false. Conversely, if (ii) is false then the ideal generated by all $n \times n$ Vandermonde determinants must be proper, and so is contained in some maximal ideal $M \lhd R$. In particular, $\det V \equiv 0 \pmod{M}$ for every $n \times n$ Vandermonde matrix V. Therefore, $\det V = 0$ for all $n \times n$ Vandermonde matrices V with entries in R/M. Because R/M is a field, every nonzero element is a unit. So $\det V$ is a unit of R/M if and only if V is generated by distinct elements. We conclude that R/M must have fewer than n distinct elements and so (i) is false.

The result (ii) $\Longrightarrow$ (iii) follows from the Cauchy-Binet formula, which gives the following expansion for $\det(LV)$ where L and V are not necessarily square but have a square product.

$$\det(LV) = \sum_{\substack{s \subseteq \{1,\ldots,N\} \\ |s| = n}} \det(L_{:,s}) \det(V_{s,:})$$

Here, the sum is taken over all subsets of $\{1, \ldots, N\}$ with n elements. The corresponding columns and rows are extracted from L and V respectively and the associated determinants are multiplied together. Since each $V_{s,:}$ is an $n \times n$ Vandermonde determinant, if L is a left-inverse for V, then $\det(LV) = \det(I) = 1$, and (ii) follows. For the converse result (ii) $\Longleftarrow$ (iii), a left-inverse can be explicitly constructed; see [20] for a proof. ∎

Lemma 14. *Suppose R is a commutative ring with identity, and H is an R-module generated by $\{H_1, \ldots, H_n\}$. Consider the following statements.*

(i) *Every residue field of R has at least n elements.*

(ii) *Suppose H is the R-module generated by some set $\{H_1, \ldots, H_n\}$. Then H is also generated by the set $\{H_1 + rH_2 + \cdots + r^{n-1}H_n \mid r \in R\}$*

Then (i) $\implies$ (ii). *If the H_i are also a basis for H, then* (i) $\iff$ (ii).

Proof. To prove (ii), it suffices to show that each H_i is a linear combination of terms of the form $H_1 + rH_1 + \cdots + r^{n-1}H_n$. If (i) holds, then by Lemma 13 there is some $N \times n$ left-invertible Vandermonde matrix $V \in R^{N \times n}$. Suppose V is generated by $r_1, \ldots, r_N$, and suppose $L \in R^{n \times N}$ is a left-inverse of V. Then it is straightforward to check that

$$H_i = \sum_{j=1}^{N} L_{ij}\left(H_1 + r_j H_2 + \cdots + r_j^{n-1}H_n\right) \qquad \text{for } i = 1, \ldots, n \tag{6}$$

as required. For the converse, suppose (ii) holds. Then an equation of the form (6) must hold for some $L \in R^{n \times N}$ and $r_1, \ldots, r_N \in R$. If the H_i form a basis for H, then the coefficients corresponding to each H_i in (6) must vanish. Therefore,

$$\sum_{j=1}^{N} L_{ij} r_j^{k-1} = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases} \qquad \text{for } i, k = 1, \ldots, n$$

In other words, $LV = I$, where V is the (left-invertible) Vandermonde matrix generated by $r_1, \ldots, r_N$. ∎

Lemma 14 may be specialized to polynomials and thus yields a sufficient condition under which a $f(r) = 0$ for all $r \in R$ implies that f is the zero polynomial.

Corollary 15. *Suppose R is a commutative ring with identity and every residue field of R has at least n elements. Suppose $f \in R[x]$ and $\deg f \leq n-1$. If $f(r) = 0$ for all $r \in R$, then $f = 0$.*

Proof. Suppose $f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$. Applying Lemma 14 to the module $H \subseteq R$ generated by $\{a_0, \ldots, a_{n-1}\}$, we conclude that H is generated by $\{f(r) \mid r \in R\}$. But $f(r) = 0$ for all $r \in R$, therefore $H = \{0\}$, and so $a_0 = \cdots = a_{n-1} = 0$. ∎

We now have the tools we need to prove our main invariance result for rings.

Proof of Theorem 6. The result is trivial if $m = 1$ or $n = 1$. In this case, either GK or KG is scalar, so every S is QI with respect to every G. Further, $K \operatorname{adj}(I_R - GK) = \operatorname{adj}(I_R - KG)K = K$, so the right-hand side always holds as well. We assume from now on that $m, n \geq 2$.

Suppose S is QI with respect to G, and let $K \in S$. Using Proposition 5, write:

$$K \operatorname{adj}(I_R - GK) = -\sum_{i=1}^{m} p_i K(I_R - GK)^{i-1}$$

$$= \sum_{i=1}^{m} h_i K(GK)^{i-1}$$

where the $h_i \in R$ are obtained by expanding each $(I_R - GK)^{i-1}$ term and collecting like powers of GK. If $2_R \in U(R)$, then by Lemma 11 all terms in the sum are in S. Since S is an R-module, it follows that $K \operatorname{adj}(I_R - GK) \in S$.

13

Conversely, suppose that S is not QI with respect to G. Therefore, there exists some $K_0 \in S$ such that $K_0 G K_0 \notin S$. We proceed by way of contradiction. Suppose that $K \operatorname{adj}(I_R - GK) \in S$ for all $K \in S$. In particular, it must hold for $K = r K_0$ with $r \in R$. Therefore, we conclude that

$$r K_0 \operatorname{adj}(I_R - r G K_0) \in S \qquad \text{for all } r \in R$$

Because each entry of the adjugate matrix is the determinant of an $(m-1) \times (m-1)$ minor, it follows that $\operatorname{adj}(I_R - r G K_0)$ is a polynomial in r of degree at most $m-1$ with coefficients in $R^{m \times m}$. Letting $\operatorname{adj}(I_R - r G K_0) = \sum_{i=0}^{m-1} r^i B_i$, we obtain

$$(K_0 B_0) r + (K_0 B_1) r^2 + \cdots + (K_0 B_{m-1}) r^m \in S \qquad \text{for all } r \in R$$

If every residue field of R has at least $m+1$ elements, we conclude from Lemma 14 that $K_0 B_i \in S$ for $i = 0, \ldots, m-1$. Recall the identity (2), and let $A = I_R - x G K_0 \in M_n(R[x])$. We then obtain

$$(I_R - x G K_0)(B_0 + x B_1 + \cdots + x^{m-1} B_{m-1}) = \det(I_R - x G K_0) I_R$$
$$= (q_0 + q_1 x + \cdots + q_m x^m) I_R$$

where the $q_i \in R$ only depend on G and K_0. Because of the ring isomorphism $M_m(R[x]) \cong (M_m(R))[x]$, the equation above is a polynomial identity in $(M_m(R))[x]$. So we may collect like powers of x and set all coefficients to zero. For the first two coefficients, we have $B_0 = q_0 I$ and $B_1 - G K_0 B_0 = q_1 I_R$. Note also that $B_0 = I_R$, which follows because $B_0 = \operatorname{adj}(I_R) = I_R$. Therefore, $B_1 = G K_0 + q_1 I_R$. Based on our earlier conclusion that $K_0 B_i \in S$, we have $K_0 B_1 = K_0 G K_0 + q_1 K_0 \in S$. Since $K_0 \in S$ by assumption and S is an R-module, it follows that $K_0 G K_0 \in S$, a contradiction. If we use the identity $K \operatorname{adj}(I - GK) = \operatorname{adj}(I - KG) K$, we now have $\operatorname{adj}(I - KG) \in \mathbb{R}^{n \times n}$. Carrying out a similar argument, we deduce that the result also holds when every residue field of R has at least $n+1$ elements, thus completing the proof. ∎

The counterexample given in Section IV is an example for which $2_R \notin U(R)$. In that case, Theorem 6 fails because Lemma 11 fails. One way to avoid the technical requirement that $2_R \in U(R)$ is to strengthen the notion of quadratic invariance. For example, if we require that $K_1 G K_2 \in S$ for all $K_1, K_2 \in S$ then Lemma 11 follows without the requirement that $2_R \in U(R)$. Thus, we would obtain a weaker version of the first part of Theorem 6. The proof of Theorem 7 uses similar machinery to that used in the proof of Theorem 6.

Proof of Theorem 7. As in Theorem 6, the cases $m = 1$ and $n = 1$ are trivial, so we assume $m, n \geq 2$. If $\operatorname{char} \mathbb{F} \neq 2$, then $2_\mathbb{F} \neq 0$, which means $2_\mathbb{F}$ is invertible. Furthermore, $\mathbb{F}$ contains at least $\min(m,n) + 1$ distinct elements, so we may apply both parts of Theorem 6. Therefore, S being quadratically invariant with respect to G is equivalent to $K \operatorname{adj}(I - GK) \in S$ for all $K \in S$. If we restrict to $K \in S \cap M$, then $\det(I - GK) \neq 0$. So we deduce that

$$h(K) = -\big(\det(I - GK)\big)^{-1} K \operatorname{adj}(I - GK) \in S$$

Since $h : M \to M$, it follows that $h(S \cap M) \subseteq S \cap M$, and by the involutive property of h, we deduce that $h(S \cap M) = S \cap M$.

Conversely, suppose that $h(S \cap M) = S \cap M$. Then it follows that $K \operatorname{adj}(I - GK) \in S$ for all $K \in S \cap M$. Suppose S is not quadratically invariant with respect to G, so there is some $K_0 \in S$ such that $K_0 G K_0 \notin S$. Proceeding as in the proof of Theorem 6, let $\operatorname{adj}(I_R - r G K_0) = \sum_{i=0}^{m-1} r^i B_i$ and obtain

$$(K_0 B_0) r + (K_0 B_1) r^2 + \cdots + (K_0 B_{m-1}) r^m \in S \qquad \text{for all } r \in R \text{ such that } r K_0 \in M$$

In other words, the inclusion holds whenever $\det(I - rGK_0) \neq 0$. This is a polynomial of degree m so it has at most m roots. Therefore, the constraint that $\det(I - rGK_0) \neq 0$ excludes at most m elements of $\mathbb{F}$. Therefore, $\mathbb{F}$ must contain at least $2m + 1$ elements so that there are at least $m + 1$ elements remaining after all roots of $\det(I - rGK_0)$ have been excluded. As in the proof of Theorem 6, we may apply a similar argument to conclude that $\mathbb{F}$ contains at least $2n + 1$ elements, and the proof in complete. ∎

VII-B Invariance for rationals

Rational functions are fundamentally a field, so our first invariance result follows immediately from our invariance result for fields.

Proof of Theorem 8. Note that $\mathbb{R}(\mathbf{s})$ is a field, so S is a subspace over $\mathbb{R}(\mathbf{s})$. Furthermore, $\mathbb{R}(\mathbf{s})$ has infinitely many elements and $2 \neq 0$, so the result follows from Theorem 7. ∎

The rationals become a ring when we impose a properness constraint. Furthermore, the strictly proper rationals are an ideal $\mathbb{R}(s)_{\mathrm{sp}} \lhd \mathbb{R}(s)_{\mathrm{p}}$. We can also check that this ideal is maximal, because the proper rationals that are not strictly proper are precisely the units of $\mathbb{R}(s)_{\mathrm{p}}$. Indeed, $\mathbb{R}(s)_{\mathrm{sp}}$ is the unique maximal ideal of $\mathbb{R}_{\mathrm{p}}$, so $\mathbb{R}_{\mathrm{p}}$ is a *local ring*.

We may use these structural facts to prove the following lemma, which gives a condition that guarantees the invertibility of $I - GK$.

Lemma 16. *Suppose* $G \in \mathbb{R}(s)_{\mathrm{sp}}^{m \times n}$ *and* $K \in \mathbb{R}(s)_{\mathrm{p}}^{n \times m}$. *Then* $(I - GK)$ *is invertible, and* $(I - GK)^{-1} \in \mathbb{R}(s)_{\mathrm{p}}^{m \times m}$.

Proof. Since $\mathbb{R}(s)_{\mathrm{sp}} \lhd \mathbb{R}(s)_{\mathrm{p}}$ is maximal and $G \in \mathbb{R}(s)_{\mathrm{sp}}^{m \times n}$, we have

$$\det(I - GK) \equiv \det(I) \equiv 1 \pmod{\mathbb{R}(s)_{\mathrm{sp}}}$$

It follows that $\det(I - GK) \notin \mathbb{R}(s)_{\mathrm{sp}}$, and is therefore a unit. So $I - GK$ is invertible. ∎

The motivation for choosing a strictly proper G and a proper K is inspired by classical feedback control. If we think of the proper rationals as transfer functions, a strictly proper K means that the controller has no direct feedthrough term, a common assumption that ensures well-posedness of the closed-loop interconnection. We now present the proof of the invariance result for proper rationals.

Proof of Theorem 9. Note that $\mathbb{R}(s)_{\mathrm{p}}$ is a ring and $2 \in \mathbb{R}(s)_{\mathrm{p}}$. Because $\mathbb{R}(s)_{\mathrm{p}}$ has a unique maximal ideal $\mathbb{R}(s)_{\mathrm{sp}}$, the only residue field is $\mathbb{R}(s)_{\mathrm{p}} / \mathbb{R}(s)_{\mathrm{sp}} = \mathbb{R}$, which has infinitely many elements. Applying Theorem 6, we conclude that S is QI with respect to G if and only if $K \operatorname{adj}(I - GK) \in S$ for all $K \in S$. However, $I - GK$ is always invertible by Lemma 16, so $K \operatorname{adj}(I - GK) \in S$ if and only if $h(K) \in S$. By the involutive property of h, this is equivalent to $h(S) = S$. ∎

Proof of Theorem 10. First, note that $\mathbb{R}(\mathbf{x}, \mathbf{s})_{\mathrm{p}} \cong \big(\mathbb{R}(\mathbf{x})\big)(\mathbf{s})_{\mathrm{p}}$, so we may think of the multivariate transfer functions as proper transfer functions in $s_1, \ldots, s_k$ with coefficients that are rational functions of $x_1, \ldots, x_\ell$. As in Theorem 9, we still have $2 \in \big(\mathbb{R}(\mathbf{x})\big)(\mathbf{s})_{\mathrm{p}}$, but there is no longer a unique maximal ideal. Indeed, the maximal ideals are the sets:

$$M_i = \big\{ f \in \big(\mathbb{R}(\mathbf{x})\big)(\mathbf{s})_{\mathrm{p}} \mid f \text{ is strictly proper in } s_i \big\}$$
$$= \Big(\big(\mathbb{R}(\mathbf{x})\big)(\mathbf{s} \setminus s_i)_{\mathrm{p}} \Big)(s_i)_{\mathrm{sp}}$$

and it is easy to check that the corresponding residue fields $\big(\mathbb{R}(\mathbf{x})\big)(\mathbf{s})_{\mathrm{p}} / M_i$ each have infinitely many elements. Furthermore, $\det(I - GK) \equiv \det(I) \equiv 1 \pmod{M_i}$ for each i, as in the proof of Lemma 16. So $\det(I - GK)$ does not belong to any maximum ideal and must therefore be a unit. The rest follows as in the proof of Theorem 9. ∎

15

VIII   Summary

In this paper, we give an algebraic treatment of quadratic invariance, the well-known condition under which decentralized control synthesis may be reduced to a convex optimization problem. Our results hold for commutative rings with identity, and in particular specialize to the natural system-theoretic case of proper rational functions in one variable, as well as multidimensional rational functions. This formulation has the advantage of avoiding some of the technicalities in analytic treatments. In particular, notions of topology, limits, or norms are not required. Thus, quadratic invariance may be viewed as a purely algebraic concept.

Acknowledgments

References

[1] B. Bamieh, F. Paganini, and M. Dahleh. Distributed control of spatially invariant systems. *IEEE Transactions on Automatic Control*, 47(7):1091–1107, 2002.

[2] C. Beck. On formal power series representations for uncertain systems. *IEEE Transactions on Automatic Control*, 46(2):314–319, 2001.

[3] C. L. Beck, J. Doyle, and K. Glover. Model-reduction of multidimensional and uncertain systems. *IEEE Transactions on Automatic Control*, 41(10):1466–1477, 1996.

[4] V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, 2000.

[5] M. L. Curtis. *Abstract linear algebra*. Springer, 1990.

[6] R. D'Andrea and G. E. Dullerud. Distributed control design for spatially interconnected systems. *IEEE Transactions on Automatic Control*, 48(9):1478–1495, 2003.

[7] M. Fliess, M. Lamnabhi, and F. Lamnabhi-Lagarrigue. An algebraic approach to nonlinear functional expansions. *IEEE Transactions on Circuits and Systems*, 30(8):554–570, 1983.

[8] E. Fornasini and G. Marchesini. Doubly-indexed dynamical systems: State-space models and structural properties. *Theory of Computing Systems*, 12(1):59–72, 1978.

[9] Y.-C. Ho and K.-C. Chu. Team decision theory and information structures in optimal control problems—Part I. *IEEE Transactions on Automatic Control*, 17(1):15–22, 1972.

[10] R. E. Kalman. Algebraic structure of linear dynamical systems, I. the module of Σ. *Proceedings of the National Academy of Sciences of the United States of America*, 54(6):1503–1508, 1965.

[11] D. Klarner. Algebraic theory for difference and differential equations. *The American Mathematical Monthly*, 76(4):366–373, 1969.

[12] A. Lamperski and L. Lessard. Optimal state-feedback control under sparsity and delay constraints. In *3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pages 204–209, 2012.

[13] S. Lang. *Undergraduate algebra*. Springer, 2005.

[14] C. Langbort, R. S. Chandra, and R. D'Andrea. Distributed control design for systems interconnected over an arbitrary graph. *IEEE Transactions on Automatic Control*, 49(9):1502–1519, 2004.

[15] L. Lessard. *Tractability of complex control systems*. PhD thesis, Stanford University, 2011.

[16] L. Lessard and S. Lall. An algebraic framework for quadratic invariance. In *IEEE Conference on Decision and Control*, pages 2698–2703, 2010.

[17] L. Lessard and S. Lall. A state-space solution to the two-player decentralized optimal control problem. In *Allerton Conference on Communication, Control, and Computing*, pages 1559–1564, 2011.

[18] X. Qi, M. V. Salapaka, P. G. Voulgaris, and M. Khammash. Structured optimal and robust control with multiple criteria: a convex solution. *IEEE Transactions on Automatic Control*, 49(10):1623–1640, 2004.

[19] F. Riesz and B. Sz.-Nagy. *Functional analysis*. Dover Publications, 1990.

[20] D. W. Robinson. The classical adjoint. *Linear Algebra and its Applications*, 411:254–276, 2005.

[21] R. Roesser. A discrete state-space model for linear image processing. *IEEE Transactions on Automatic Control*, 20(1):1–10, 1975.

[22] M. Rotkowitz, R. Cogill, and S. Lall. Convexity of optimal control over networks with delays and arbitrary topology. *International Journal of Systems, Control and Communication*, 2(1):30–54, 2010.

[23] M. Rotkowitz and S. Lall. Decentralized control information structures preserved under feedback. In *IEEE Conference on Decision and Control*, pages 569–575, 2002.

[24] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, 51(2):274–286, 2006.

[25] Ş. Sabău and N. Martins. Necessary and sufficient conditions for stabilizability subject to quadratic invariance. In *IEEE Conference on Decision and Control*, pages 2459–2466, 2011.

[26] P. Shah and P. A. Parrilo. $\mathcal{H}_2$-optimal decentralized control over posets: A state-space solution for state-feedback. *IEEE Transactions on Automatic Control*, 58(12):3084–3096, 2013.

[27] H. Shin and S. Lall. Decentralized control via Gröbner bases and variable elimination. *IEEE Transactions on Automatic Control*, 57(4):1030–1035, 2012.

[28] M. C. Smith. On stabilization and the existence of coprime factorizations. *IEEE Transactions on Automatic Control*, 34(9):1005–1007, 1989.

[29] M. Vidyasagar. *Control system synthesis: a factorization approach*. MIT Press, 1985.

[30] N. Wiener. *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*. MIT press, 1964.

[31] H. S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1):131–147, 1968.

[32] D. Youla, H. Jabr, and J. Bongiorno Jr. Modern Wiener-Hopf design of optimal controllers–Part II: The multivariable case. *IEEE Transactions on Automatic Control*, 21(3):319–338, 1976.