Distributed Control of Spatially Invariant Systems

Bassam Bamieh, *Member, IEEE*, Fernando Paganini, *Member, IEEE*, and Munther A. Dahleh, *Fellow, IEEE*

***Abstract*—We consider distributed parameter systems where the underlying dynamics are spatially invariant, and where the controls and measurements are spatially distributed. These systems arise in many applications such as the control of vehicular platoons, flow control, microelectromechanical systems (MEMS), smart structures, and systems described by partial differential equations with constant coefficients and distributed controls and measurements. For fully actuated distributed control problems involving quadratic criteria such as linear quadratic regulator (LQR), $\mathcal{H}_2$ and $\mathcal{H}_\infty$, optimal controllers can be obtained by solving a parameterized family of standard finite-dimensional problems. We show that optimal controllers have an inherent degree of decentralization, and this provides a practical distributed controller architecture. We also prove a general result that applies to partially distributed control and a variety of performance criteria, stating that optimal controllers inherit the spatial invariance structure of the plant. Connections of this work to that on systems over rings, and systems with dynamical symmetries are discussed.**

***Index Terms*—Distributed control, infinite-dimensional systems, optimal control, robust control, spatially invariant systems.**

I. Introduction

THE VAST majority of automatic control implementations to this day are spatially "lumped," in the sense that the controller interfaces with the physical system at a fixed and relatively small number of actuators and sensors. Therefore the spatially distributed aspect of the dynamics is only internal to the system, usually modeled as an infinite-dimensional *state*; the theory of *distributed parameter systems* (see [1]–[4]) has focused on methods (C_0-semigroups, operator equations, etc.) for the precise mathematical treatment of these internal dynamics, which is significantly more difficult than the finite-dimensional theory. The constructive aspects of these theories typically deal with insuring that finite-dimensional approximation schemes converge. Problems with spatially distributed sensing and actuation pose an additional challenge, in that finite dimensional approximations always produce systems with large scale inputs and outputs. With few exceptions, such problems have not been thoroughly studied due to the perception of their technological infeasibility.

Recently, however, technological progress is bringing dramatic changes to this picture. In particular, advances in microelectromechanical systems (MEMS) make feasible the idea of microscopic devices with actuating, sensing, computing, and telecommunications capabilities. Distributing a large array of such devices in a spatial configuration gives unprecedented capabilities for control; some recent examples are distributed flow control [5]–[8] for drag reduction, and "smart" mechanical structures [9]–[11]. At a larger scale, the control of networks of autonomous units is gaining attention, e.g., in infinite strings of vehicles (recently known as Platoons) [12]–[14], and in cross directional (CD) control [15]–[17] in the chemical process industry. For these applications the control variables are also distributed in space, in addition to the internal states. While, of course, practical systems will involve a finite number of sensors and actuators, the correct abstraction in the limit for a large array is that of a spatio-temporal system [18], where all variables are indexed in space and time. Important questions that arise are i) how to design controllers for these systems with regard to global objectives; and ii) how can these control algorithms be implemented in a distributed array.

In this paper, we study these questions for an important class of problems. The key property which we exploit is *spatial invariance*, which means that there is a notion of translation in some spatial coordinates, with respect to which the plant dynamics are invariant. The systems we consider and the optimal controllers we design in this paper are typically *infinite dimensional*. However, for the spatially invariant class, we show that quadratically optimal controllers can be designed by solving parameterized families of finite-dimensional problems (via parameterized families of matrix Riccati equations). This is an exact solution of the infinite dimensional problem. We also show that the optimal infinite-dimensional controllers have an inherently semidecentralized architecture (which we refer to as "localized"). This architecture consists of a distributed infinite array of finite dimensional controllers with separation structure, and observer and state feedback operators which are spatial convolutions. One of our main results is that the corresponding convolution kernels (which determine the communication requirements in the controller array) have exponential rates of decay spatially. These facts allow us to argue for *spatial truncation* to implement these controllers rather than the standard practice of *modal truncation*. In contrast to the latter, the former preserves the inherently localized structure of optimal controllers, and has exponential rather than polynomial rates of approximation convergence.

Manuscript received May 29, 1998; revised December 10, 1999, August 5, 2000, February 20, 2001, and April 10, 2001. Recommended by Associate Editor H. Ozbay. The work of B. Bamieh was supported by the National Science Foundation under CAREER award ECS-96-24152, and by the Air Force Office of Scientific Research-PRET 442530-25380. The work of F. Paganini was supported by the National Science Foundation under CAREER award ECS-9875056. The work of M. A. Dahleh was supported by the National Science Foundation under award ECS-9612558 and by the Air Force Office of Scientific Research under Grant F49620-95-0219.

B. Bamieh is with the Department of Mechanical and Environmental Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: bamieh@engineering.ucsb.edu).

F. Paganini is with the Department of Electrical Engineering, University of California at Los Angeles, Los Angeles, CA 90024 USA (e-mail: paganini@seas.ucla.edu).

M. A. Dahleh is with Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: dahleh@lids.mit.edu).

Publisher Item Identifier 10.1109/TAC.2002.800646.

The notion of spatial invariance should be viewed as the counterpart to time invariance for spatio-temporal systems. Another main result in this paper is that for spatially invariant plants, one can restrict attention to spatially invariant controllers without any loss in performance. This fact provides a significant simplification of controller implementation and design, and further generalizes our results on quadratic problems. We prove this invariance property for a large class of systems and induced-norm robust control measures.

Our work is related to several earlier results in the literature, and we briefly mention here some connections. The earliest use of a spatial-invariance concept for control design is the work of [12] for infinite strings of systems, with application to vehicle arrays. This work and subsequent extensions in [13], [19] already include the use of transform methods over $\mathbb{Z}$ for control optimization. The work on discretized partial differential equations [20] can be viewed as similar to our results for the group $\mathbb{Z}_n$. This case has also appeared elsewhere in the literature in the context of block-circulant transfer function matrices [16], and is related to the "lifting" technique for N-periodic systems [21]. In our work we generalize these notions to spatial coordinates over arbitrary groups, including continuous domains. For the case of infinite domains, we study the inherently localized structure of optimal controllers which to our knowledge has not been previously addressed in the theory. Interestingly however, this fact seems to have been observed in special examples earlier. In pursuing the question of sensor location [22], [23], it was observed that optimal feedback kernels (referred to as "functional gains") are inherently localized.

The notion of spatio-temporal systems has also received some attention in the early 1980s in the literature on systems over rings [18], [24]–[26], where conditions for stability and stabilizability were obtained in terms of a parameterized families of finite dimensional conditions. Again, our work here generalizes these notions and we furthermore show that these techniques can also effectively solve optimal control problems.

Another line of related research is the work on symmetries of linear dynamical systems [27], [28], where among other results, it was shown that plants with certain dynamical symmetries can be stabilized by controllers with the same symmetry (i.e., without "symmetry breaking" in the terminology of [28]). In this paper we consider systems with similar spatial symmetries and show that one can also achieve optimal performance without breaking symmetry.

More recently and in parallel to our work, distributed sensing and actuation problems have also been addressed using linear matrix inequality (LMI) techniques [29], [30], where the distributed nature of inputs and outputs and limited communication requirements are incorporated in the theory.

Our presentation is organized as follows. In Section II, we introduce the formalism of systems over spatial groups to study such invariances, and collect some facts from generalized commutative Fourier analysis. The simplest case (which we label *fully actuated*) is when all spatial coordinates are of this nature, and the actuation/sensing is fully distributed across them. As discussed in Section III, these systems can be studied by a Fourier transform over the spatial domain, which *block-diagonalizes* the dynamics into a family of finite dimensional systems over spatial frequency. In particular, optimal control design with respect to quadratic performance criteria (LQR, $\mathcal{H}_2$, $\mathcal{H}_\infty$) can be decoupled into a family of standard finite dimensional problems over spatial frequency, as shown in Section IV.

In Section V, we analyze the structure of these optimal controllers, with particular attention to their localization or decentralization in space. This is a key issue for the implementation of these optimal strategies in a distributed array. In this regard, we show in Section V-A that a natural distributed architecture can be provided, where local controllers observe the local state and communicate with neighboring actuators and sensors. In Section V-B, we study the spatial decay rates of the related convolution kernels, which directly affects the communication requirements for the array. By analytically extending solutions of parameterized Riccati equations, we prove that the decay rates are spatially exponential.

In Section VI, we move to the general situation where spatially invariant coordinates exist but do not fully parameterize the plant dynamics; in this case diagonalization can still be performed, but to yield a finite-dimensional problem it must be combined with some form of lumped approximation in the remaining coordinates. In these cases our techniques still provide a computational reduction in controller design, as well as an important insight into the architecture of distributed controllers. We end the paper with a general result which shows that under a variety of (not necessarily quadratic) performance criteria, the optimal controller can always be taken to have the same invariances as the underlying system, i.e., in the terminology of [28], one can achieve optimality without symmetry breaking. Conclusions are given in Section VII.

II. Mathematical Preliminaries

In this section, we introduce a formalism to study the question of *spatial invariance* in spatially distributed dynamical systems. In full generality, spatial invariance can be defined whenever a *group* of symmetries acts on the spatial coordinates, and the dynamics commute with this group. In this paper, we will restrict ourselves to the special case when some spatial variables themselves form a group, and the symmetries considered are translations in this group.

A. Groups, Translations and Invariance

In the sequel, $\mathbb{G}$ will denote a locally compact, abelian (LCA) *group*, see [31]. Special cases of this include

1) $\mathbb{G} = \mathbb{R}$;
2) $\mathbb{G} = \partial\mathbb{D}$ (unit circle);
3) $\mathbb{G} = \mathbb{Z}$ (integers);
4) $\mathbb{G} = \mathbb{Z}_n$, (finite group of integers modulo n).

In addition, one can consider direct products of such spaces, $\mathbb{G} := \mathbb{G}_1 \times \cdots \times \mathbb{G}_d$, e.g., $\mathbb{R}^d$, $\mathbb{Z}^d$, or the cylinder $\partial\mathbb{D} \times \mathbb{R}$. Such examples cover all the cases of interest in this paper, but still the abstract formulation is convenient to treat all cases at once. In this paper, the group $\mathbb{G}$ will represent a number of spatial coordinates, each varying in one of the above groups. There may also be additional coordinates in the problem; in other words, signals will have the form $u(x, \xi, t)$, where $x = (x_1, \ldots, x_d)$ varies in a group $\mathbb{G} := \mathbb{G}_1 \times \cdots \times \mathbb{G}_d$, ξ contains additional spatial coordinates varying in some *set*, and t is (discrete or continuous) time.

The group operation (denoted by $+$, in the case of $\partial\mathbb{D}$ this corresponds to addition of arcs) introduces a translation map $x \mapsto x + x_o$ on $\mathbb{G}$, and a translation operator for functions on $\mathbb{G}$ in the natural way: $(T_{x_o} f)(x) := f(x - x_o)$. There is a natural *measure* on $\mathbb{G}$ which is invariant under translations, and positive over open sets; this is called the Haar measure and is unique up to normalization. (e.g., Lebesgue measure in $\mathbb{R}$ or $\partial\mathbb{D}$, counting measure on $\mathbb{Z}$ or $\mathbb{Z}_n$). We denote it by dx. We will be mainly concerned with complex functions on $\mathbb{G}$ which are square integrable with respect to the Haar measure

$$\mathcal{L}_2^n(\mathbb{G}) := \left\{ f : \mathbb{G} \rightarrow \mathbb{C}^n \| \|f\|_2^2 := \int_{\mathbb{G}} |f(x)|^2 dx < \infty \right\}. \quad (1)$$

An operator $A : \mathcal{D}(A) \rightarrow \mathcal{L}_2^m(\mathbb{G})$ with domain $\mathcal{D}(A) \subset \mathcal{L}_2^n(\mathbb{G})$ is said to be *translation invariant* if $T_x : \mathcal{D}(A) \rightarrow \mathcal{D}(A)$ and $AT_x = T_x A$ for every translation T_x. Two important examples are as follows.

- For $\mathbb{G} = \mathbb{R}$, the differentiation operator

$$A : f(x) \mapsto \frac{\partial f}{\partial x}. \quad (2)$$

  The domain of this operator is the set of functions in $\mathcal{L}_2(\mathbb{R})$ with derivative in $\mathcal{L}_2(\mathbb{R})$; it is clear that A is translation invariant. Note that A is unbounded, but it is a closed operator with dense domain (see, e.g., [2]); due to their importance in partial differential equation models, the theory of distributed parameter systems is usually developed for this larger class, see [3]. More generally, a PDE operator whose spatial domain is a group, and has constant coefficients is translation invariant.
- A spatial convolution: let H_x be a family of operators indexed over $x \in \mathbb{G}$, and define

$$\mathbf{H} : f \mapsto \int_{\mathbb{G}} H_{x-\zeta} f(\zeta) d\zeta \quad (3)$$

  where the integral corresponds to the Haar measure. Under appropriate assumptions, this operator is well defined and spatially invariant.

B. Fourier Analysis on Groups

One of the main advantages of the spatial invariance property over a group, is that Fourier transforms can be introduced to *diagonalize* the relevant operators, in the same way as time invariance is exploited in standard system theory. This section contains a brief overview of Fourier analysis over groups; for a full account see [31].

General Fourier analysis (also known as commutative harmonic analysis) consists on mapping functions on $\mathbb{G}$ to functions on a *dual group* $\hat{\mathbb{G}}$; in full generality, $\hat{\mathbb{G}}$ can be identified with the set of homomorphisms from $\mathbb{G}$ to $\partial\mathbb{D}$ (the so-called characters); for our purposes it suffices to consider Table I, where the entries correspond to the Fourier transform, Fourier series, Z-transform, and discrete Fourier transform, respectively.

TABLE I
COMMUTATIVE GROUPS AND THEIR DUALS

$\mathbb{G}$	$\mathbb{R}$	$\partial\mathbb{D}$	$\mathbb{Z}$	$\mathbb{Z}_p$
$\hat{\mathbb{G}}$	$\mathbb{R}$	$\mathbb{Z}$	$\partial\mathbb{D}$	$\mathbb{Z}_p$

The Fourier transform $\mathcal{F}$ associates a function $\{f(x)\}$ on $\mathbb{G}$ with a function $\{\hat{f}(\lambda)\}$ on $\hat{\mathbb{G}}$ (for a general definition see [31]). A few properties are

1) $\mathcal{F}$ is linear;
2) $\mathcal{F}$ transforms convolutions into products;
3) with appropriate normalizations in the measures dx and $d\lambda$, $\mathcal{F}$ is an isometric isomorphism from $\mathcal{L}_2(\mathbb{G})$ to $\mathcal{L}_2(\hat{\mathbb{G}})$; in particular the Plancherel theorem states that

$$\|f\|_2^2 = \int_{\mathbb{G}} |f(x)|^2 dx = \int_{\Lambda} |\hat{f}(\lambda)|^2 d\lambda = \|\hat{f}\|_2^2. \quad (4)$$

The last property above implies that one can identify the spaces $\mathcal{L}_2(\mathbb{G})$ and $\mathcal{L}_2(\hat{\mathbb{G}})$; in particular, every operator A on a dense domain in $\mathcal{L}_2(\mathbb{G})$ is identified with an operator $\hat{A} = \mathcal{F}A\mathcal{F}^{-1}$ on a dense domain in $\mathcal{L}_2(\hat{\mathbb{G}})$. The main advantage of this identification in that translation invariant operators are associated with *multiplication* operators in the transformed domain.

Definition 1: A multiplication operator $\mathcal{D}(\hat{A}) \subset \mathcal{L}_2^n(\hat{\mathbb{G}})$ is defined by

$$[\hat{A}\hat{f}](\lambda) := \hat{A}(\lambda)\hat{f}(\lambda)$$

almost everywhere, where $\{\hat{A}(\lambda)\} : \hat{\mathbb{G}} \longrightarrow \mathbb{C}^{m \times n}$ is a measurable matrix-valued function.

The matrix-valued function $\{\hat{A}(\lambda)\}$ is called the *symbol* of the operator $\hat{A}$. With a slight abuse of notation, we use the same letter $\hat{A}$ to denote both the operator and its symbol. We will also denote such symbols by either $\hat{A}(\lambda)$ or $\hat{A}_\lambda$ to make formulas more readable. It is easily seen that if $\hat{A}$ is a multiplication operator, the corresponding A on $\mathcal{L}_2(G)$ is translation invariant; in what follows, we restrict the attention to translation invariant operators with this property.

Assumption 1: We will consider translation invariant operators $A : \mathcal{D}(A) \rightarrow \mathcal{L}_2^m(\mathbb{G})$ such that the corresponding $\hat{A} : \mathcal{D}(\hat{A}) \rightarrow \mathcal{L}_2^m(\hat{\mathbb{G}})$, $\mathcal{D}(\hat{A}) \subset \mathcal{L}_2^n(\hat{\mathbb{G}})$ is a multiplication operator, and the function $\hat{A}(\lambda)$ is continuous.

This assumption is general enough to include spatial convolutions with kernels in $\mathcal{L}_1(\mathbb{G})$, as well as the differentiation operators which appear in constant coefficient PDE models. As is well known [32], the induced norm of multiplication operators is given by the $\mathcal{L}_\infty$ norm of the defining function, in other words

$$\|A\| = \|\hat{A}\| = \sup_{\lambda \in \hat{\mathbb{G}}} \sigma_{\max}\left(\hat{A}(\lambda)\right)$$

and the finiteness of the right-hand side is a necessary and sufficient condition for the boundedness of the operator A.

III. SPATIAL INVARIANCE WITH FULL ACTUATION: STATE-SPACE MODELS

This paper is concerned with spatio–temporal systems where the relevant signals are indexed by a spatial coordinate in addition to time. In this section, we focus on a special class of control problems, where the following are true.

1) All the spatial coordinates, denoted collectively by x, vary in a group $\mathbb{G}$. Modulo this coordinate, the dynamics are finite dimensional.
2) The system actuators and sensors are fully distributed over this coordinate.
3) The dynamics are spatially invariant with respect to translations in this coordinate.

To be more precise, we adopt a state-space description for this class of systems. The first two assumptions imply that if $\mathbf{\Psi}$, $\mathcal{U}$ and $\mathcal{Y}$ denote the state, input, and output spaces respectively, $\dim\{\mathbf{\Psi}/\mathbb{G}\}$, $\dim\{\mathcal{U}/\mathbb{G}\}$, $\dim\{\mathcal{Y}/\mathbb{G}\}$, are all finite. In other words they can be expressed as vector-valued functions $\psi(x,t)$, $u(x,t)$ and $y(x,t)$. A general linear model is, thus, of the form (in continuous time)

$$\frac{\partial}{\partial t}\psi(x,t) = [A\psi](x,t) + [Bu](x,t) \tag{5}$$

$$y(x,t) = [C\psi](x,t) + [Du](x,t). \tag{6}$$

Definition 2: The system (5) and (6) is called *spatially invariant* if the operators A, B, C, D are translation invariant and satisfy Assumption 1.

The entire treatment will be in terms of $\mathcal{L}_2$ signal spaces: in (5) and (6), at a fixed instant of time the signals ψ, u and y are assumed to be elements of $\mathcal{L}_2^n(\mathbb{G})$ (respectively, $\mathcal{L}_2^p(\mathbb{G})$ and $\mathcal{L}_2^r(\mathbb{G})$) for some finite vector dimensions n, p, r. A, B, C, D are translation invariant operators between spaces $\mathcal{L}_2(\mathbb{G})$ of appropriate dimensions, and are static (no dependence on t). The third assumption above means that these operators are translation invariant, in the sense defined in Section II. For example, A, B, C, D could be matrices whose elements are PDE operators (in x) with constant coefficients, spatial shift operators, spatial convolution operators, or a linear combination of several such operators.

In general, some of the operators will be unbounded, so the notion of a solution to (5) requires some care, and involves the theory of C_0 semigroups of operators; for a full discussion, see [3].

A. Examples

A standard example of a spatially invariant system is the heat equation with fully distributed control over either an infinite domain ($\mathbb{G} = \mathbb{R}$) or with periodic boundary conditions ($\mathbb{G} = \partial\mathbb{D}$).

An example where the spatial domain is discrete is longitudinal control and string stability of vehicular platoons ($\mathbb{G} = \mathbb{Z}$). Such problems have been considered using transform techniques where the property of spatial invariance was utilized [12], [13]. If each vehicle in an infinite string is modeled as a moving mass with second-order dynamics (with normalized coefficients), we obtain

$$\begin{bmatrix}\dot{p}_i\\ \dot{v}_i\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}p_i\\ v_i\end{bmatrix} + \begin{bmatrix}0\\ u_i - u_{i-1}\end{bmatrix} + \begin{bmatrix}0\\ w_i - w_{i-1}\end{bmatrix}$$

for all $i \in \mathbb{Z}$, and where p_i, v_i are the relative position and velocity errors between the i'th and $(i-1)$ vehicle, respectively. u_i, w_i are the control and disturbance inputs into the ith vehicle. In such a problem, it is typically desired to regulate errors down to zero, and this can be captured by quadratic performance objective such as

$$J = \sum_{i\in\mathbb{Z}} \int_0^\infty \left(\alpha_1 p_i^2(t) + \alpha_2 v_i^2(t) + \alpha_3 u_i^2(t)\right) dt.$$

Note that in the above model, the A operator is obviously spatially invariant. The B operator is actually $[0 \quad I - T_{-1}]^T$, where T_{-1} is the operator of translation by -1. And as usual, we can define a regulated output whose $\mathcal{L}_2$ norm is the quadratic objective defined above. All the system operators thus defined will be spatially invariant. One can then pose distributed $\mathcal{H}_2$ and $\mathcal{H}_\infty$ problems for such systems. We finally note here that the stability of the distributed system implies the so-called "string stability" of the vehicular platoon.

Other examples, were the spatial variables are of a mixed discrete and continuous nature can be accommodated in our framework.

B. Block-Diagonalization

The main observation we will exploit in the next few sections is that by taking a Fourier transform, the system (5) and (6) is *diagonalized* into the decoupled form

$$\frac{d}{dt}\hat{\psi}(\lambda,t) = \hat{A}(\lambda)\hat{\psi}(\lambda,t) + \hat{B}(\lambda)\hat{u}(\lambda,t) \tag{7}$$

$$\hat{y}(\lambda,t) = \hat{C}(\lambda)\hat{\psi}(\lambda,t) + \hat{D}(\lambda)\hat{u}(\lambda,t) \tag{8}$$

where $\hat{A}(\lambda)$, $\hat{B}(\lambda)$, $\hat{C}(\lambda)$, $\hat{D}(\lambda)$ are multiplication operators in accordance with assumption 1. Now, the transformed system (7) and (8) is in effect a decoupled family of standard finite-dimensional linear time-invariant (LTI) systems over the frequency parameter λ.

Stability: The decoupling achieved by Fourier transformation allows for standard techniques from finite dimensional system theory to be applied to these distributed problems. Consider first the question of stability of the autonomous equation

$$\frac{\partial}{\partial t}\psi = A\psi \tag{9}$$

with $\psi \in \mathcal{L}_2(\mathbb{G})$. Definitions of stability, asymptotic stability and exponential stability have been studied for such systems (see [2]), which extend, with some complications, the finite dimensional theory. In the translation invariant case, this question can be studied by means of the diagonalized system $(d/dt)\hat{\psi} = \hat{A}\hat{\psi}$.

We will assume that the operator A generates a strongly continuous (C^o) semigroup[2] on $\mathcal{L}_2(\mathbb{G})$, which we refer to as $\{T(t)\}$.

Definition 3: The system (9) is *exponentially stable* if

$$\|T(t)\| \le Me^{-\alpha t}, \qquad t \ge 0$$

for some $M, \alpha > 0$.

Note that the corresponding semigroup $\hat{T}(t) := \mathcal{F}T(t)\mathcal{F}^{-1}$ on $\mathcal{L}_2(\hat{\mathbb{G}})$ is a semigroup of multiplication operators with symbol $e^{t\hat{A}(\lambda)}$. Thus, the stability condition can be recast as

$$\sup_{\lambda \in \hat{\mathbb{G}}} \sigma_{\max}\left(e^{t\hat{A}(\lambda)}\right) \leq Me^{-\alpha t}. \tag{10}$$

This implies that checking exponential stability is *almost* equivalent to checking "pointwise" the stability of the decoupled systems. The precise statement is in the following theorem.

Theorem 1: If A is the generator of a strongly continuous semigroup, then the following two statements about the system (9) are equivalent:

1) the system is exponentially stable;
2) for each $\lambda \in \hat{\mathbb{G}}$, $\hat{A}_\lambda$ is stable, and the solution of the family of matrix Lyapunov equations

$$\hat{A}_\lambda^* P_\lambda + P_\lambda \hat{A}_\lambda = -I \tag{11}$$

is bounded, i.e., $\sup_{\lambda\in\hat{\mathbb{G}}} \|P_\lambda\| < \infty$.

Proof: The main point is that $\sup_{\lambda\in\hat{\mathbb{G}}} \|P_\lambda\| < \infty$ is equivalent to the boundedness of the multiplication operator on $\mathcal{L}_2(\hat{\mathbb{G}})$ defined by the symbol $\{P_\lambda\}$.

$2 \Rightarrow 1$: This follows [33, Th. 5.1.3]. The solution P is a bounded Hermitian operator that satisfies the weak version of the operator Lyapunov equation, and thus, $V(\psi) :=< P\psi, \psi >$ is a Lyapunov function for the system [33, Th. 5.1.3].

$1 \Rightarrow 2$: Since the semigroup $\hat{T}(t)$ is made up of multiplication operators, the boundedness condition (10) is

$$\sup_{\lambda\in\hat{\mathbb{G}}} \lambda_{\max}\left(e^{t\hat{A}_\lambda^*}\, e^{t\hat{A}_\lambda}\right) \leq M^2 e^{-2\alpha t}.$$

Therefore, if we define P_λ to be the solution of (11) for each λ, we can bound

$$\begin{aligned}\|P_\lambda\| &= \int_0^\infty \lambda_{\max}\left(e^{t\hat{A}_\lambda^*} e^{t\hat{A}_\lambda}\right) dt \\ &\leq \int_0^\infty M^2 e^{-2\alpha t} dt < \infty\end{aligned}$$

independently of λ. ∎

The aforementioned theorem then implies that checking exponential stability can be handled by finite dimensional tools plus a search over λ.

Stabilizability: Similar statements can be made about the question of stabilizability of the system

$$\frac{\partial}{\partial t}\psi = A\psi + Bu \tag{12}$$

where B is a bounded operator on $\mathcal{L}_2(\mathbb{G})$. The following statements can be generalized to the case where B is unbounded, but we state the bounded case here for simplicity.

Definition 4: The system (12) is *exponentially stabilizable* if there exists an operator $F : \mathcal{L}_2(\mathbb{G}) \longrightarrow \mathcal{L}_2(\mathbb{G})$ such that $A - BF$ generates an exponentially stable C^o semigroup on $\mathcal{L}_2(\mathbb{G})$.

It turns out that checking stabilizability can be done by a pointwise solution to a parameterized family of finite dimensional Riccati equations.

Theorem 2: Let A be the generator of a C^o semigroup, and B be bounded. Then, the system in (12) is exponentially stabilizable if and only if the following two conditions hold:

1) for all $\lambda \in \hat{\mathbb{G}}$, the pair $(\hat{A}_\lambda, \hat{B}_\lambda)$ is stabilizable;
2) the solution of the family of matrix Riccati equations

$$\hat{A}_\lambda^* P_\lambda + P_\lambda \hat{A}_\lambda - P_\lambda \hat{B}_\lambda \hat{B}_\lambda^* P_\lambda + I = 0 \tag{13}$$

is bounded, i.e., $\sup_{\lambda\in\hat{\mathbb{G}}} \|P_\lambda\| < \infty$.

Proof: We first note that condition 1) implies that for all λ there exists a positive definite stabilizing solution to the algebraic Riccati equation (ARE) (13), and we require condition 2) to insure that this results in a bounded operator on $\mathcal{L}_2(\hat{\mathbb{G}})$.

Sufficiency: This is clear since $V(\psi) := \langle P\psi, \psi\rangle$ is a Lyapunov function for the closed loop system with feedback $-B^*P$.

Necessity: This is essentially in [3, Th. 4.1.8], which concludes that there must exist a unique bounded Hermitian solution to a weak operator version of (13). In this case, that solution gives the value function $V(\psi) = \langle P\psi, \psi\rangle$ to the related LQR problem. The value function is clearly spatially invariant, and therefore the corresponding Hermitian operator P is spatially invariant. Thus, the operator ARE collapses to (13), together with the boundedness condition. ∎

We remark here that the above two theorems imply that checking stability or stabilizability of spatially invariant systems can be done by checking the same condition for the finite-dimensional decoupled systems for every frequency $\lambda \in \hat{\mathbb{G}}$.

Remark: There are examples where it is necessary to check the boundedness conditions in Theorems 1 and 2. However, these appear to be mathematical constructions rather than physical examples. At large spatial frequencies (i.e., as $\lambda \to \infty$) the dominant mechanism in physical systems is dissipation. In other words, for large λ, physical systems are stable and thus stabilizable, and the conditions are automatically satisfied in the limit $\lambda \to \infty$.

The boundedness condition is certainly not needed when the group $\hat{G}$ is compact (i.e., for the case of spatially discrete systems $\mathbb{G} = \mathbb{Z}, \mathbb{Z}_n$), since it follows from the continuity of solutions. The above results are then very similar to results on so-called spatio-temporal systems from the theory of systems over rings [18], [25], [26]. We summarize this in the following corollary.

Corollary 3: If the group $\hat{\mathbb{G}}$ is compact, then

1) the system in (9) is exponentially stable if and only if for every $\lambda \in \hat{\mathbb{G}}$, the matrix $\hat{A}_\lambda$ is stable;
2) the system in (12) is exponentially stabilizable if and only if for every $\lambda \in \hat{\mathbb{G}}$, the matrix pair $(\hat{A}_\lambda, \hat{B}_\lambda)$ is stabilizable.

We note here that in the compact case, no technical assumptions on the operators A, B are necessary other than the continuity of the matrix-valued functions $\{\hat{A}_\lambda\}$, $\{\hat{B}_\lambda\}$, which then imply that A, B are bounded operators on $\mathcal{L}_2(\mathbb{G})$.

IV. Optimal Control With Quadratic Measures

We now discuss briefly the possible settings for optimal and robust control problems for such systems. There are two possible lines of investigation. The first would be to combine the spatial Fourier transform with a Laplace transform over time to obtain a multidimensional system transfer function. To illustrate, assume that (9) is exponentially stable, then $(sI - A)$ has a bounded inverse for $\mathrm{Re}(s) \geq 0$, and we can define

$$H_\lambda(s) = \hat{C}_\lambda(sI - \hat{A}_\lambda)^{-1}\hat{B}_\lambda + \hat{D}_\lambda \tag{14}$$

which is a $p \times r$ matrix-valued function on $\hat{\mathbb{G}} \times \{\mathrm{Re}(s) \geq 0\}$, continuous in λ and analytic in s. Then analysis and synthesis problems for such systems become equivalent problems for a certain class of multidimensional systems. In this setting, causality (or stability) for both plants and controllers is only relevant with respect to the variable s, and not the spatial transform variable λ. This is a crucial difference between these problems and other multidimensional systems problems.

In this paper, we will take another approach which is more expedient for quadratic problems. We transform only the spatial variables, as was done in (7) and (8). The main observation is that since $\mathcal{L}_2$ norms are preserved by the Fourier transform, any optimal control problem on (5) and (6) involving *quadratic* signal norms (e.g., LQR, $\mathcal{H}_2$ or $\mathcal{H}_\infty$ problems, see below) will be equivalent to an analogous problem for (7) and (8). Thus, the original distributed problem is converted to a parameterized family of finite dimensional state space problems.

A. The Distributed LQR

We begin by studying the distributed LQR problem. There is an abundant amount of literature (see [3], [34]) on this problem, characterizing the optimum in terms of a solution to an operator Riccati equation, in an analogous fashion to the finite dimensional theory. Such equations are difficult to solve in general, but for the class of spatially invariant systems, the problem diagonalizes exactly into a parameterized family of finite dimensional LQR problems. This is now explained; for simplicity only the infinite horizon problem is discussed.

Consider the problem of minimizing the functional

$$J = \int_0^\infty \langle Q\psi, \psi\rangle + \langle Ru, u\rangle dt \tag{15}$$

subject to the dynamics (5), and $\psi(x, 0) = \psi_0(x) \in \mathcal{L}_2(\mathbb{G})$. Our main assumption is that A, B, Q, and R are translation invariant operators (further assumptions are listed in Theorem 4).

Taking spatial transforms, utilizing the facts that translation invariant operators transform to multiplication operators, and that inner products are preserved, the problem can then be rewritten as the minimization of

$$J = \int_{\hat{\mathbb{G}}} \int_0^\infty \left(\hat{\psi}^*\lambda(t)\hat{Q}\lambda\hat{\psi}\lambda(t) + \hat{u}_\lambda^*(t)\hat{R}_\lambda\hat{u}\lambda(t)\right) dt d\lambda \tag{16}$$

subject to (7) and $\hat{\psi}(\lambda, 0) = \hat{\psi}_0(\lambda)$, and where $\hat{Q}$ and $\hat{R}$ are the Fourier symbols of the operators Q and R, respectively. Now it is clear from (16) and (7) that the problem decouples over λ, that is, it is "block-diagonal" with the blocks parameterized by λ. At a fixed λ, it amounts to no more than a classical finite-dimensional LQR problem. Therefore, the unique solution to this problem is achieved by the translation invariant state feedback $u = -R^{-1}B^*Px$, where P is a translation invariant operator whose Fourier symbol $\hat{P}(\lambda)$ is the positive–definite solution to the parameter-dependent matrix ARE

$$\hat{A}_\lambda^*\hat{P}_\lambda + \hat{P}_\lambda\hat{A}_\lambda + -\hat{P}_\lambda\hat{B}_\lambda\hat{R}_\lambda^{-1}\hat{B}_\lambda^*\hat{P}_\lambda + \hat{Q}_\lambda = 0 \tag{17}$$

for all $\lambda \in \hat{\mathbb{G}}$.

The main observation here is that when A, B, Q, R are translation invariant operators, then the solution to the operator ARE in the LQR problem is also a translation invariant operator. The exact conditions under which this yields a stabilizing controller are:

Theorem 4: Consider the LQR problem (15), (5), where A, B, Q, R are translation invariant operators, with $R > 0, Q \geq 0$. If (A, B) and $(A^*, Q^{1/2})$ are exponentially stabilizable, then

1) the solution to the family of matrix ARE's in (17) is uniformly bounded, i.e., $\sup_{\lambda\in\hat{\mathbb{G}}} P(\lambda) < \infty$;
2) the translation invariant feedback operator $K = -R^{-1}B^*P$ is exponentially stabilizing.

Proof: As in the proof of Theorem 2, this follows from [3, Cor. 4.17 and Th. 4.18] after noting that the solution of the operator ARE must be a spatially invariant operator. This follows from the spatial invariance of the value function of the LQR problem. This last fact is clear: if A, B, Q, R are spatially invariant operators, then for any initial state ψ_o, we have from the LQR problem definition that $J(\psi_o) = J(T_x\psi_o)$, where T_x is any translation. ■

We should note here that the above is a generalization of the notion of solving certain LQR problems by so-called modal decomposition. The important difference is that in this paper we do not advocate the standard "modal truncation" as the method to implement a finite-dimensional approximation of the optimal infinite-dimensional controller. Rather, we will proceed further and analyze the properties of the resulting feedback operators. This will yield a more natural approximation scheme that we refer to as "spatial truncation." These issues will be studied in Section V-B.

B. $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Control

In this section, we briefly describe how the preceding methodology applies also to two disturbance rejection problems of the standard form given in Fig. 1. The generalized plant G is a linear, space/time invariant distributed system, which admits a state-space representation as in (5) and (6). The feedback K, which is also distributed, must internally (exponentially) stabilize the system and minimize a certain norm of the closed loop map T_{zw}. As will be shown in more generality in Section VI, no performance loss occurs by restricting the design to controllers which are themselves space/time invariant. Under these circumstances, the closed loop $H := T_{zw}$ is a space/time invariant system, which can be represented by either the convolution

$$z(x, t) = \int_G \int_0^\infty h(\zeta, \tau)w(x - \zeta, t - \tau)d\zeta d\tau \tag{18}$$

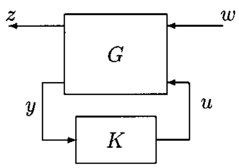


Fig. 1. The standard problem.

(in the continuous time case) or the transfer function representation

$$\hat{z}(\lambda, s) = \hat{H}(\lambda, s)\hat{w}(\lambda, s). \tag{19}$$

The stabilizing property of K is required to guarantee (18) and (19) are meaningful under suitable signal classes. For our purposes, we will require that $\hat{H}(\lambda, s)$ be bounded, and analytic in the second variable over $\hat{\mathbb{G}} \times \{\mathrm{Re}(s) > 0\}$.

This in particular, implies that H is a well-defined operator on $\mathcal{L}_2(\mathbb{G} \times \mathbb{R})$, with induced norm

$$\|H\|_{\mathcal{L}_2 \to \mathcal{L}_2} = \|H\|_\infty := \sup_{\lambda \in \hat{\mathbb{G}}, \omega \in \mathbb{R}} \overline{\sigma}(\hat{H}(\lambda, j\omega)) \tag{20}$$

which we call the $\mathcal{H}_\infty$ norm of the system; note, however, that the Hardy space (analytic) structure refers only to the second variable; from the point of view of the spatial frequency this is only an $\mathcal{L}_\infty$ norm. Thus, the $\mathcal{H}_\infty$ control problem in this context is to find a stabilizing K which minimizes the previous quantity.

Another system specification which translates naturally to the frequency domain is the $\mathcal{H}_2$ criterion, given by

$$\|H\|_2^2 = \frac{1}{2\pi} \int_{\hat{\mathbb{G}}} \int_{-\infty}^{\infty} \mathrm{tr}\left(\hat{H}^* \lambda(j\omega) \hat{H} \lambda(j\omega)\right) d\lambda d\omega. \tag{21}$$

As in the finite dimensional case, this norm can be used to measure the response of the system to stochastic disturbances. Once again, we remark that the Hardy space structure is with respect to the temporal-frequency variables alone.

The distributed $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control design problems are defined in terms of the following state-space description of the plant G in Fig. 1:

$$\frac{\partial}{\partial t}\psi(.,t) = A\psi(.,t) + B_1 w(.,t) + B_2 u(.,t) \tag{22}$$
$$z(.,t) = C_1\psi(.,t) + D_{12}u(.,t) \tag{23}$$
$$y(.,t) = C_2\psi(.,t) + D_{21}w(.,t) \tag{24}$$

here A, B_1, B_2, C_1, C_2, D_{21}, D_{12} are translation invariant operators, and it is assumed that (A, B_2) and (A^*, C_2^*) are exponentially stabilizable. Note that these conditions can be tested by a parameterized family of finite dimensional stabilizability checks as in Theorem 2.

The control design problems consist of finding controllers of the form

$$\frac{\partial}{\partial t}\psi_K(.,t) = A_K\psi_K(.,t) + B_K y(.,t) \tag{25}$$
$$u(.,t) = C_K\psi_K(.,t) + D_K y(.,t) \tag{26}$$

where the state is a real separable Hilbert space, and A_K is the generator of a C_o semigroup, such as the the closed-loop system be exponentially stable, and either of

- $\|T_{zw}\|_2$ is minimized. ($\mathcal{H}_2$ control);
- $\|T_{zw}\|_\infty < 1$ ($\mathcal{H}_\infty$ control).

These problems have been addressed in the literature for general classes of distributed parameter systems [3], [34]). The solutions generally available are in terms of the solvability of two *operator* AREs. As we now indicate, spatial invariance allows us to solve these operator AREs as a parameterized family of matrix AREs. This can be observed by looking at the Fourier transformed version of (22)–(24)

$$\frac{d}{dt}\hat{\psi}_\lambda(t) = \hat{A}_\lambda\hat{\psi}_\lambda(t) + \hat{B}_{1,\lambda}\hat{w}_\lambda(t) + \hat{B}_{2,\lambda}\hat{u}_\lambda(t) \tag{27}$$
$$\hat{z}_\lambda(t) = \hat{C}_{1,\lambda}\hat{\psi}_\lambda(t) + \hat{D}_{12,\lambda}\hat{u}_\lambda(t) \tag{28}$$
$$\hat{y}_\lambda(t) = \hat{C}_{2,\lambda}\hat{\psi}_\lambda(t) + \hat{D}_{21,\lambda}\hat{w}_\lambda(t). \tag{29}$$

The system is thus reduced to a parameterized family of finite dimensional LTI systems over $\lambda \in \hat{\mathbb{G}}$. Now for both the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ problems, the optimization can be decoupled over frequency since quadratic performance measures are preserved by Fourier transformation

- from (21), the distributed $\mathcal{H}_2$ problem amounts to solving a family of standard $\mathcal{H}_2$ problems over λ, then integrating for the overall cost;
- the $\mathcal{H}_\infty$ feasibility question $\|T_{zw}\|_\infty < 1$ can be imposed as a family of standard $\mathcal{H}_\infty$ conditions $\|T_{zw}(\lambda, \cdot)\|_\infty < 1$.

This means that the standard finite dimensional theory can apply at every λ, and the optimal controllers can be found by applying the standard finite dimensional Riccati equations as in [35] at each λ. The only additional technical requirement is to show boundedness over λ of the resulting solutions; for this purpose some regularity conditions must be imposed, which specialize those of [34] to the spatially invariant case. For the sake of brevity we omit the detailed statements; the interested reader is referred to [36].

As a consequence of the above, we find that the optimal $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers for spatial invariant plants are themselves spatially invariant; a generalization of this fact will be discussed in Section VI.

V. The Structure of Quadratically Optimal Controllers

The optimal controllers obtained in Section IV have the following attractive properties:

- they provide *global* performance guarantees. In particular, they will ensure overall stability;
- they can be effectively computed by a family of low dimensional problems across spatial frequency.

However, we have not considered the issue of *implementation* of the control algorithm. In this regard, rather than a highly complex *centralized* controller with information from all the distributed array, it would be desirable to have distributed intelligence, where each actuator runs a local algorithm with information from the neighboring sensors. In this section, we analyze the optimal schemes from this perspective. Relevant questions are as follows.

- (Section V-A) Does the control law lend itself to a distributed architecture?

- (Section V-B) To what degree is information from far away sensors required? Notice that this pertains to approximate diagonalization in the *original* spatial variables.

A. Local Controller Architecture

We now illustrate the surprisingly intuitive and appealing architecture of quadratically optimal controllers (by this we mean the LQR, $\mathcal{H}_2$ and the central $\mathcal{H}_\infty$ controller).

First note that since the ARE solutions for all three problems are translation-invariant operators, then their controllers are spatially invariant systems; in particular, the same algorithm must be run at each actuator location, the influence of each sensor depending on its position relative to the actuator.

To understand the structure of this algorithm, let us examine more closely the optimal $\mathcal{H}_2$ controller; it has the following realization:

$$\begin{aligned}\frac{\partial}{\partial t}\psi_K(x,t) =& A\psi_K(x,t) + B_2 u(x,t)\\ &+ L \star (C_2\psi(x,t) - C_2\psi_K(x,t)) \quad (30)\\ u(x,t) =& F \star \psi_K(x,t) \quad (31)\end{aligned}$$

where the state feedback and estimator "gains" are $F := -B_2^* P_1$ and $L := P_2 C_2^*$, where P_1 and P_2 are the solutions to the operator Riccati equations, themselves spatially invariant. Thus in the original spatial coordinates, F and L are spatial convolutions, which is previously emphasized by the $\star$ notation.

The above implies the following structure of the optimal controller.

a) A *distributed estimator* whose local state is $\psi_K(x,t)$. Note that to propagate this state [(30)], one needs to know the outputs of neighboring estimators, and convolve the prediction errors with the kernel of L (the size of this neighborhood is determined by the spread of L). We note that at a given $x_o \in \mathbb{G}$, the local controller state $\psi_K(x_o,t)$ has a physical interpretation; it is the estimate of the system's local state $\psi(x_o,t)$.

b) The feedback at position x is given by $u(x,t)$ which is computed by convolving neighboring state estimates with the kernel of F (the size of this neighborhood is determined by the spread of F).

Thus the optimal laws are directly amenable to a distributed implementation, with localized actuation and information passing. What determines the degree of localization, and thus the communication burden for the array, is the spread of the convolution operators L and F. Note that the open loop system operators A, B_2 and C_2 are typically PDE operators, therefore localized[1].

However, in general the Riccati solutions P_1 and P_2 will not be differential operators ($\hat{P}_1(\lambda)$ and $\hat{P}_2(\lambda)$ are not rational in general, see the example in Section V-B), and their convolution kernels will have a spread, reflecting the need of information passing within the array.

In Section V-B, we will provide means of evaluating the spread of P_1 and P_2; in particular we will see that these convolution kernels decay exponentially in space; thus the optimal control laws have an inherent degree of decentralization. From a practical perspective, the convolution kernels can be spatially truncated to form "local" convolution kernels that have performance close to the optimal, and preserve the appealing architecture described earlier.

[1]For open-loop operators of the form (3), their convolution kernels would also contribute to the controller spread.

B. The Degree of Spatial Localization

We will study the localization issue for systems with unbounded spatial domains; for these we can ask the question of how the controller gains decay as we move away in space. For concreteness, we focus on the case $\mathbb{G} = \mathbb{R}$ (and $\hat{\mathbb{G}} = \mathbb{R}$), analogous ideas apply to the discrete case $\mathbb{G} = \mathbb{Z}$. We first consider the simplest example of LQR optimization to illustrate the spatial localization properties.

Example: Consider the heat equation on an infinite bar with distributed heat injection

$$\frac{\partial}{\partial t}\psi(x,t) = c\frac{\partial^2}{\partial x^2}\psi(x,t) + u(x,t). \quad (32)$$

Here the group $\mathbb{G}$ is the real line. The standard Fourier transform yields the transformed system

$$\frac{d}{dt}\hat{\psi}(\lambda,t) = -c\lambda^2\hat{\psi}(\lambda,t) + \hat{u}(\lambda,t). \quad (33)$$

Taking, for example, $Q = qI$ (multiple of the identity) and $R = I$ in the LQR cost (15), the corresponding (scalar) parameterized Riccati equation is

$$-2c\lambda^2\,\hat{p}(\lambda) - \hat{p}^2(\lambda) + q = 0 \quad (34)$$

which has the positive solution

$$\hat{p}(\lambda) = -c\lambda^2 + \sqrt{c^2\lambda^4 + q}. \quad (35)$$

In the transform domain, the optimal control will be of the form $\hat{u}(\lambda,t) = \hat{k}(\lambda)\hat{\psi}(\lambda,t)$, with $\hat{k}(\lambda) = -\hat{p}(\lambda)$. Note that even though the system and the cost are rational in λ, the optimal control is irrational. This in particular implies that it cannot be implemented by a "completely localized" PDE in x and t. Indeed the control law takes the spatial convolution form

$$u(x,t) = \int_{\mathbb{R}} k(x-\zeta)\psi(\zeta)d\zeta$$

where the convolution kernel $k(x) = -p(x)$, and $p(x)$ is the inverse Fourier transform of $\hat{p}(\lambda)$. In this particular case, noticing that

$$\hat{p}(\lambda) = \frac{q}{c\lambda^2 + \sqrt{c^2\lambda^4 + q}} \in L_1(\mathbb{R})$$

we see that $p(x)$ is a continuous function of x. The degree of spatial decentralization is characterized by the "spread" of $p(x)$ as a function of x. We now show that this kernel has some degree of inherent "localization."

If we measure this spread in terms of the exponential decay of $p(x)$, then we can bring in Laplace transform tools to relate this decay to the existence of an *analytic continuation* of the Fourier transform $\hat{p}(\lambda)$ into a vertical strip of the complex plane.

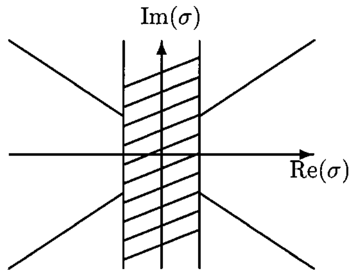


Fig. 2. Analytic continuation region for $\hat{p}(\lambda) = -c\lambda^2 + \sqrt{c^2\lambda^4 + q}$.

In particular, $\hat{p}(\lambda)$ can be extended to the function

$$\hat{p}_e(\sigma) = c\sigma^2 + \sqrt{c^2\sigma^4 + q}$$

of $\sigma \in \mathbb{C}$, and such that $\hat{p}(\lambda) = \hat{p}_e(j\lambda)$. We note that $\hat{p}_e$ is analytic in a region of the complex plane which avoids the four branch cuts shown by the diagonal lines in Fig. 2. Thus, the Fourier transform of $\hat{p}$ can be analytically extended to the strip

$$\left\{\sigma \in \mathbb{C} : |\mathrm{Re}(\sigma)| < \frac{\sqrt{2}}{2}\left(\frac{q}{c^2}\right)^{1/4}\right\}.$$

Now, a consequence of Theorem 5 is that $p(x)$ must decay exponentially, more precisely

$$|p(x)|e^{\eta|x|} \overset{|x|\to\infty}{\longrightarrow} 0, \text{ for } 0 < \eta < \frac{\sqrt{2}}{2}\left(\frac{q}{c^2}\right)^{1/4}.$$

Since $\{k(x)\}$ decays exponentially with $|x|$, it can be truncated to form a "localized" feedback convolution operator whose closed loop performance is close to the optimal.

Remark: We note that in this particular problem, an interesting tradeoff seems to be in place: in the limit of "cheap" control (i.e., $q \longrightarrow \infty$) the analyticity region grows and the controller becomes more decentralized (approximates a purely local feedback). It seems thus possible that there is an inherent tradeoff between actuator authority and controller centralization, in the sense that low authority actuators need feedback from more distant sensors to achieve optimality.

In the remainder of this section we generalize these ideas beyond this specific example. We begin by precisely stating the relationship between analytic continuation and exponential decay of the inverse transform. We draw from the theory of Fourier-Laplace transforms as described in [37, Ch. 7]. Here the theory is set up in the Schwartz space of distributions $\mathcal{D}'(\mathbb{R})$, and a key role is played by the subspace of *tempered* distributions $\mathcal{S}'(\mathbb{R})$, which have a well-defined Fourier transform $\mathcal{F}$. For our purposes it suffices to note that a tempered *function* on $\mathbb{R}$ is a locally integrable function $f(x)$ that grows no faster than polynomially as $|x| \to \infty$, and tempered distributions are obtained by finite generalized derivatives of such functions; thus they may have singularities (Dirac δ's, etc.) but their growth rate is no more than polynomial.

The generalization of (bilateral) Laplace transforms works as follows.[2] For a given distribution $f \in \mathcal{D}'(\mathbb{R})$, define the set

$$\Gamma = \{\eta \in \mathbb{R} : e^{-\eta x} f \in \mathcal{S}'(\mathbb{R})\}.$$

The Laplace transform $F(\sigma)$ on the vertical strip $\Gamma + j\mathbb{R} = \{\sigma \in \mathbb{C} : \mathrm{Re}(\sigma) \in \Gamma\}$ is defined by

$$F(\eta + j\lambda) = \mathcal{F}[e^{-\eta x} f](j\lambda). \tag{36}$$

With this definition in place, we now state the following result from [37, Th. 7.4.2].

Theorem 5: Let Γ be an open interval in $\mathbb{R}$, and $F(\sigma)$ be an analytic function on the strip $\Gamma + j\mathbb{R}$, such that for every compact set $\Gamma_0 \subset \Gamma$, there exist $C, N > 0$ such that

$$|F(\sigma)| \le C(1 + |\sigma|)^N \tag{37}$$

holds for $\mathrm{Re}(\sigma) \in \Gamma_0$. Then there exists a distribution $f \in \mathcal{D}'(\mathbb{R})$ such that $e^{-\eta x} f \in \mathcal{S}'(\mathbb{R})$ for every $\eta \in \Gamma$, and satisfying (36).

If we specialize this result to $\Gamma = (-\beta, \beta)$ we see that analytic functions $F(\sigma)$ on the strip $\{\sigma \in \mathbb{C} : |\mathrm{Re}(\sigma)| < \beta\}$ which satisfy the growth bound (37), have inverse Fourier transforms f such that $e^{-\eta x} f$ is of "tempered" growth in x, for every $|\eta| < \beta$. For instance if f itself is a continuous function (as in the example above), then $f(x)$ must decay to zero as $|x| \to \infty$, faster than any exponential $e^{-|\eta||x|}$; this is the decay result we are looking for.

To generalize this kind of result to a class of optimal control problems, we must guarantee that the Riccati solutions $\hat{P}(\lambda)$ have an analytic continuation $P_e(\sigma)$ to a vertical strip of this kind, and that they satisfy a growth bound of the form (37). This will now be pursued, focusing on the LQR Riccati (17), and for simplicity setting $R = I$. We narrow the set of problems with the following assumptions.

Assumption 2:

i) The functions $\hat{A}(\lambda)$, $\hat{B}(\lambda)$, $\hat{Q}(\lambda)$ have analytic extensions $\hat{A}_e(\sigma)$, $\hat{B}_e(\sigma)$, $\hat{Q}_e(\sigma)$ to the strip $\mathcal{S} = \{\sigma \in \mathbb{C} : |\mathrm{Re}(\sigma)| < \alpha\}$, which are *rational* functions. Note that this class includes differential operators which are common in "open loop" PDE models. We use the notation $\hat{A}_e^{\sim}(\sigma) = \hat{A}_e^T(-\sigma)$, where ^{T} denotes transpose.

ii) For every $\sigma \in \mathcal{S}$, $\left(\hat{A}_e(\sigma), \hat{B}_e(\sigma)\hat{B}_e^{\sim}(\sigma)\right)$ is stabilizable.

iii) For every $\sigma \in \mathcal{S}$, if for some vectors $y, z \in \mathbb{C}^n$, $z^T\hat{B}_e(\sigma)\hat{B}_e^{\sim}(\sigma)y = 0$, then either $z^T\hat{B}_e(\sigma) = 0$, or $\hat{B}_e^T(-\sigma)y = 0$ (or both). In other words, the range space of $\hat{B}_e^T(\sigma)$ does not contain vectors orthogonal to those in the range space of $\hat{B}_e^T(-\sigma)$.

This last assumption is satisfied for examples where $\hat{B}_e(\sigma)$ is either a scalar or a vector. We remark here that these assumptions are perhaps not the most general under which the following results can be obtained. Under these assumptions, we will show

[2]This material is from [37, Sec. 7.4]: here Fourier transforms are defined on the *real* axis and analytic continuations are done in *horizontal* strips, so minor changes are needed to adapt to our current notation.

that the ARE solution $\hat{P}(\lambda)$ admits an analytic continuation to a strip $\Gamma + j\mathbb{R}$ around the imaginary axis (possibly smaller than $\mathcal{S}$) which satisfies the hypothesis of Theorem 5. This will mean that

$$\hat{K}(\lambda) = -\hat{B}^*(\lambda)\hat{P}(\lambda)$$

also satisfies these conditions, and therefore corresponds to a convolution kernel $K(x)$ such that $K(x)e^{\eta x}$ is a tempered distribution for $\eta \in \Gamma$. This means, in a distribution sense, that we have exponential decay of the controller "spread." We will also study means of computing the strip and decay bound.

Remark: The importance of this exponential decay is seen more clearly when comparing it with the decay of the Fourier transform, which is algebraic in λ. Thus, we see that at least in a qualitative sense, the spatial domain is more appropriate to perform truncations than the Fourier domain. This is an important comment, because a commonly used method for controller design in distributed parameter systems is to "pick a number of modes," i.e., truncate in the Fourier domain. This "modal truncation" ignores the localization of the controller in the spatial coordinates.

As an illustration, let K_T be the spatially truncated convolution operator defined by truncating the convolution kernel of the optimal feedback K

$$K_T(x) = \begin{cases} K(x), & |x| \le T \\ 0, & |x| > T. \end{cases}$$

Assume that $K - K_T$ is a function (i.e., any distribution components of K are supported in $[-T, T]$, this is the typical case); under this assumption, exponential decay can be expressed as $|K(x) - K_T(x)| \le Me^{-\eta|x|}$ for $|x| > T$ and thus

$$\begin{aligned} \|K - K_T\|_{\mathcal{L}_2 \to \mathcal{L}_2} &= \sup_{\lambda \in \hat{\mathbb{G}}} |\hat{K}(\lambda) - \hat{K}_T(\lambda)| \\ &\le \int_{-\infty}^{\infty} |K(x) - K_T(x)| dx \le \frac{2M}{\eta} e^{-\eta T}. \end{aligned}$$

This means that if we establish exponential decay of the kernel, we can make the operator norm of the truncation error $K - K_T$ exponentially small as T becomes large. By standard small-gain arguments, this approximation property guarantees that the truncated feedback is stabilizing and approximates the optimal performance. As an illustration, we demonstrate the small-gain stability argument. Write the *truncated* closed loop $d\psi/dt = (A + BK_T)\psi$ as the interconnection of the two systems

$$\begin{aligned} \frac{d\psi}{dt} &= (A + BK)\psi + v, \\ v &= B(K_T - K)\psi. \end{aligned} \tag{38}$$

The mapping $v \to \psi$ in (38) is bounded on $\mathcal{L}_2$ since K is exponentially stabilizing (see [33, Th. 5.1.5]); if B is a bounded operator[3], the feedback mapping $\psi \to v$ has arbitrarily small norm. The small-gain theorem now implies stability of the closed-loop system.

[3]This is not the most general situation under which a small-gain argument is possible. We assume the boundedness of B here for simplicity.

1) Analytic Continuation: The first step will be to prove the existence of an analytic continuation to the Riccati equation solution. We note that in this section, we drop the subscript $_e$ and the "hats" from all extension functions for simplicity of notation. To perform the extension, define the following generalized Hamiltonian matrix

$$H(\sigma) = \begin{bmatrix} A(\sigma) & -B(\sigma)B^{\sim}(\sigma) \\ -Q(\sigma) & -A^{\sim}(\sigma) \end{bmatrix}. \tag{39}$$

$H(\sigma)$ is analytic on $\sigma \in \mathcal{S}$; restricted to the $j\lambda$ axis it is the Hamiltonian matrix which corresponds to the ARE (17). Hence (see [38]), its stabilizing solution $P(j\lambda)$ is associated with the stable invariant subspace of $H(j\lambda)$; to extend $P(j\lambda)$ analytically amounts to extending this subspace. We note that for $\sigma \ne j\lambda$ we still use the notation $H(\sigma) \in$ dom (Ric) to mean that it has no purely imaginary eigenvalues and a stable eigensubspace complementary to the range of $[0 \quad I]^T$. In such cases we denote the unique stabilizing (not necessarily Hermitian) solution by Ric $(H(\sigma))$ (see Appendix A). The precise statement of analytic extension follows.

Theorem 6: Given the previous assumptions and H as in (39), with $H(j\lambda) \in \mathrm{dom}(\mathrm{Ric}) \forall\ \lambda \in \mathbb{R}$

1) if for some $\beta > 0$, $\beta < \alpha$, condition

$$\forall \omega \in \mathbb{R},\ |\mathrm{Re}(\sigma)| < \beta \quad \det(j\omega I - H(\sigma)) \ne 0 \tag{40}$$

holds, we can conclude that in the strip $\{\sigma \in \mathbb{C} : |\mathrm{Re}(\sigma)| < \beta\}$, $H(\sigma) \in \mathrm{dom}(\mathrm{Ric})$, and $P(\sigma) := \mathrm{Ric}(H(\sigma))$ is an analytic extension of $\hat{P}(j\lambda)$;

2) there exists a $\beta > 0$ such that the condition (40) is satisfied.

Proof:

Part 1) The condition (40) guarantees that no eigenvalue of $H(\sigma)$ crosses the imaginary axis as σ varies in the strip $\{\sigma \in \mathbb{C} : |\mathrm{Re}(\sigma)| < \beta\}$. Since $H(\sigma) \in \mathrm{dom}(\mathrm{Ric})$ for $\sigma = j\lambda$, we conclude that for any σ in the strip, $H(\sigma)$ has exactly n stable and n antistable eigenvalues. We note here that $H(\sigma)$ no longer has the standard Hamiltonian structure for $\sigma \ne j\lambda$, however, our assumptions guarantee that it has a stabilizing solution (no longer necessarily Hermitian) with the property that $P(\sigma) = P^T(-\sigma)$ (we refer the reader to Appendix A for the details of this argument). An additional argument in Appendix B implies that the function $P(\sigma)$ is differentiable.

It remains to show that the solution $P(\sigma)$ is analytic in the strip. To see this, note that $P(\sigma)$ is the stabilizing solution of the ARE

$$\begin{aligned} 0 = A^{\sim}(\sigma)P(\sigma) + P(\sigma)A(\sigma) \\ -P(\sigma)B(\sigma)B^{\sim}(\sigma)P(\sigma) + Q(\sigma) \end{aligned}$$

Differentiating this equation with respect to $\bar{\sigma}$ (the complex conjugate of σ), we obtain (suppressing the argument σ)

$$\begin{aligned} 0 = \frac{\partial A^{\sim}}{\partial \bar{\sigma}} P + A^{\sim} \frac{\partial P}{\partial \bar{\sigma}} + \frac{\partial P}{\partial \bar{\sigma}} A + P \frac{\partial A}{\partial \bar{\sigma}} - \frac{\partial P}{\partial \bar{\sigma}} BB^{\sim} P \\ -P \frac{\partial B}{\partial \bar{\sigma}} B^{\sim} P - PB \frac{\partial B^{\sim}}{\partial \bar{\sigma}} P - PBB^{\sim} \frac{\partial P}{\partial \bar{\sigma}} + \frac{\partial Q}{\partial \bar{\sigma}}. \end{aligned}$$

Using analyticity of A, B, Q, this equation reduces to

$$(A^{\sim} - PBB^{\sim})\frac{\partial P}{\partial \bar{\sigma}} + \frac{\partial P}{\partial \bar{\sigma}}(A - BB^{\sim}P) = 0$$

which is a Sylvester equation. Since $(A - BB^{\sim}P)$ and $(A^{\sim} - PBB^{\sim})$ are both stable in $\{\sigma \in \mathbb{C} : |\mathrm{Re}(\sigma)| < \beta\}$, then we have that $\partial P/\partial\bar{\sigma} = 0$ in that strip, i.e., P is analytic there.

Part 2. Let us write $\sigma =: \epsilon + j\lambda$. The rationality of H implies that (40) can be rewritten as

$$\begin{aligned} &\forall \epsilon \in (-\beta,\beta), \lambda,\omega \in \mathbb{R}, \quad \det\left(j\omega I - H(\epsilon + j\lambda)\right) \neq 0 \\ &\Updownarrow \\ &\forall \epsilon \in (-\beta,\beta), \lambda,\omega \in \mathbb{R}, \quad q(\omega,\lambda,\epsilon) \neq 0 \end{aligned} \tag{41}$$

where q is some polynomial in the real variables $\omega, \lambda, \epsilon$. The hypotheses imply that (41) is true at $\epsilon = 0$. We want to show that this is also true for $\epsilon \in (-\beta,\beta)$ for some $\beta > 0$. To do this, we appeal to the Tarski–Seidenberg quantifier elimination procedure [39], [40], where we can obtain the following equivalence:

$$\forall\, \lambda,\omega \in \mathbb{R},\ q(\omega,\lambda,\epsilon) \neq 0 \Leftrightarrow \tilde{q}_1(\epsilon) > 0, \ldots, \tilde{q}_n(\epsilon) > 0$$

for some n single-variable polynomials $\{\tilde{q}_i\}$. Clearly, if this last condition is satisfied at $\epsilon = 0$, it is satisfied in some nonzero interval $(-\beta,\beta)$. ∎

2) A Growth Bound: In order to apply Theorem 5 to the continuation $P(\sigma)$, we must show that its entries satisfy a polynomial growth bound of the form (37); for this, we exploit the fact that the function is *algebraic*.

We begin with the scalar case; a function $p(\sigma)$ of the complex variable σ is algebraic if it satisfies the polynomial identity $g(\sigma, p(\sigma)) \equiv 0$, where

$$g(\sigma, z) = \gamma_n(\sigma)z^n + \cdots + \gamma_1(\sigma)z + \gamma_0(\sigma) \tag{42}$$

is a two-variable polynomial, i.e., each $\gamma_i(\sigma)$ is a polynomial of complex coefficients; we assume that g is nontrivial in z, i.e., $n > 0$ and $\gamma_n(\sigma) \not\equiv 0$. The objective is to derive a growth bound for $p(\sigma)$. We first state the following.

Lemma 7: Given a monic polynomial $g(z) = z^n + \cdots + \gamma_1 z + \gamma_0$, then all roots satisfy

$$|z| \leq \prod_{i=0}^{n-1}(1 + n|\gamma_i|).$$

Proof: Take $|z| > \prod_{i=0}^{n-1}(1 + n|\gamma_i|)$. Then $|z| > 1$ and so for each $i = 0, \ldots, n-1$ we have

$$|z|^{n-i} \geq |z| \geq 1 + n|\gamma_i| > n|\gamma_i|.$$

This leads to

$$\left|\sum_{i=0}^{n-1}\gamma_i z^{i-n}\right| \leq \sum_{i=0}^{n-1}\left|\gamma_i z^{i-n}\right| < \sum_{i=0}^{n-1}\frac{1}{n} = 1$$

and therefore

$$g(z) = z^n\left(1 + \sum_{i=0}^{n-1}\gamma_i z^{i-n}\right) \neq 0.$$

∎

We can now derive the desired bound for algebraic functions.

Proposition 8: Let $p(\sigma)$ be an analytic function in the strip $\Gamma + j\mathbb{R}$, where $\Gamma \subset \mathbb{R}$ is open. Also assume that $p(\sigma)$ is algebraic as defined above. Then for every compact $\Gamma_0 \subset \Gamma$, there exist $C, N > 0$ such that for every $\sigma \in \Gamma_0 + j\mathbb{R}$

$$|p(\sigma)| \leq C(1 + |\sigma|)^N.$$

Proof: By hypothesis, $g(\sigma, p(\sigma)) \equiv 0$ where $g(\sigma, z)$ is as in (42). Since $\gamma_n(\sigma)$ is a nonzero polynomial, the set $\{\sigma \in \Gamma_0 : |\gamma_n(\sigma)| \leq 1\}$ is compact; since $p(\sigma)$ is continuous, by choosing C appropriately we can ensure that the bound holds over this set.

It remains to consider the behavior of $p(\sigma)$ for σ's where $|\gamma_n(\sigma)| > 1$. For each of these, $p(\sigma)$ is a root of the polynomial

$$z^n + \cdots + \frac{\gamma_1(\sigma)}{\gamma_n(\sigma)}z + \frac{\gamma_0(\sigma)}{\gamma_n(\sigma)}$$

so the previous lemma gives

$$\begin{aligned} |p(\sigma)| &\leq \prod_{i=0}^{n-1}\left(1 + n\left|\frac{\gamma_i(\sigma)}{\gamma_n(\sigma)}\right|\right) \leq \prod_{i=0}^{n-1}(1 + n|\gamma_i(\sigma)|) \\ &\leq C(1 + |\sigma|)^N \end{aligned}$$

for N equal to the sum of the degrees of the $\gamma_i(\sigma)$, and an appropriately chosen C. ∎

The previous results can be immediately applied to the scalar algebraic Riccati equation

$$-p^2 b(\sigma)b(-\sigma) + [a(-\sigma) + a(\sigma)]p + q(\sigma) = 0$$

under Assumption 2 . The function $p(\sigma)$ obtained from Theorem 6 is analytic on a strip, and also algebraic. Therefore, it can be bounded as required for Theorem 5.

In the matrix case, the previous results will apply provided we can go from the matrix ARE to a scalar algebraic equation for each entry of $P(\sigma)$; for details see Appendix B.

3) Computational Test for the Decay Rate: We have seen that the exponential decay rate is dictated by the width β of the strip in which we have analytic continuation, and that condition (40) provides a (possibly conservative) estimate of β. We now show that this determinant condition can be efficiently tested. Assuming again $H(\sigma)$ is a rational function, we can write the "descriptor realization"

$$H(\sigma) = D_H + C_H(\sigma E - A_H)^{-1}B_H \tag{43}$$

well defined over $\mathcal{S}$, and reduce condition (40) to

$$\det\begin{bmatrix} \sigma E - A_H & B_H \\ C_H & j\omega I - D_H \end{bmatrix} \neq 0 \tag{44}$$

for all $\omega \in \mathbb{R}$, $|\mathrm{Re}(\sigma)| < \beta$. This procedure casts the Hamiltonian condition for β as a well-posedness condition, of an analogous nature to the structured singular value theory [41]. With this inspiration, we can "close the other loop" and define

$$M(j\omega) := A_H + B_H(j\omega I - D_H)^{-1}C_H$$

which leads to the equivalent condition

$$\det(\sigma E - M(j\omega)) \neq 0, \quad \omega \in \mathbb{R}, \ |\mathrm{Re}(\sigma)| < \beta. \tag{45}$$

Note that conditions (43) and (45) are equivalent to (44) via Schur complements. So we conclude that β satisfies (40) if and only if

$$\inf_{\omega \in \mathbb{R}} |\mathrm{Re}\,[eig(E, M(j\omega))]| \geq \beta. \tag{46}$$

This last condition involves only a generalized eigenvalue computation plus a one dimensional search over the parameter ω; in general, we can plot the eigenvalue condition over a grid of frequencies, and from there estimate a suitable value of β. Invoking the preceding theory, a β satisfying (46) bounds the exponential decay of the optimal convolution kernel $P(x)$.

VI. Partially Actuated Systems: Extensions and Limitations

The main assumption in our work thus far has been that the system is fully actuated, i.e., that the controls and states are defined over the same index set $\mathbb{G}$. More precisely, we assumed that $\dim\{\mathbf{\Psi}/\mathbb{G}\}$, $\dim\{\mathcal{U}/\mathbb{G}\}$, $\dim\{\mathcal{Y}/\mathbb{G}\}$, are all finite. In this section, we first show how for systems in which $\dim\{\mathbf{\Psi}/\mathcal{U}\}$ is infinite, our techniques still provide some reduction in the complexity of the control design problem. We then prove a general result valid for nonquadratic performance measures, that optimal controllers are spatially invariant if the system is.

We first consider state models of the form

$$\frac{\partial}{\partial t}\psi(x,\xi,t) = A\psi(x,\xi,t) + Bu(x,t) \tag{47}$$

$$y(x,t) = C\psi(x,\xi,t) + Du(x,t) \tag{48}$$

with $x \in \mathbb{G}$, a group, and $\xi \in \mathcal{S}$ a set of indices with no a priori defined structure. The operators A, B, C, D are assumed to commute only with translations T_x, for $x \in \mathbb{G}$.

A more general and abstract model would be to simply regard the state as taking values in an infinite-dimensional space for every $x \in \mathbb{G}$ (this amounts to "suppressing" the variable ξ). We can then write the model as previously

$$\frac{\partial}{\partial t}\psi(x,t) = A\psi(x,t) + Bu(x,t)$$
$$y(x,t) = C\psi(x,t) + Du(x,t)$$

where now for every $x \in \mathbb{G}$ the system is infinite-dimensional with finite-dimensional inputs and outputs.

The techniques we have presented for fully actuated systems can now be applied to yield parameterized families of control problems involving *infinite*-dimensional systems. To solve each such problem, one must resort to some approximation technique. However, our development implies that *one needs to approximate only in the dimensions in which the problem is not spatially invariant.*

We now briefly present an example to illustrate this point.

A. Stabilization of Fluid Flow in a Channel

This is a problem that has attracted much recent attention [7], [8]. An example of this problem is depicted in Fig. 3, where

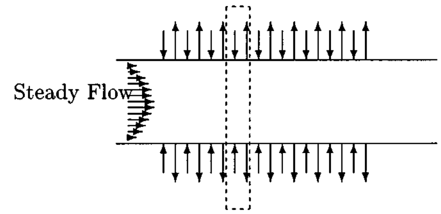


Fig. 3. 2-D Poiseuille flow with active boundary control.

the flow of a fluid in a two-dimensional (2-D) channel is being controlled by blowing/suction actuators (indicated by arrows toward the channel), and an array of flow shear sensors are used as measurement outputs (indicated by arrows away from channel). The control objective in this problem is to stabilize the nominal laminar flow. This set up is a prototype for the important problem of skin-friction drag reduction by microactuators and sensors [7], [8].

The system model for this problem can be taken to be the linearized Navier–Stokes equations about the nominal flow. If we divide the system into an array of "cells" as shown. The cells consist of vertical segments (each to include one pair of sensors and actuators as the basic cell). The state $\bar{\psi}_i$ of each cell would then represent the flow field in each vertical segment. The spatial invariance of the system with respect to a discrete index $i \in \mathbb{Z}$ representing horizontal shifts from one cell to the next can be established (assuming the underlying PDE to have coefficients constant in the horizontal direction, which happens in this case). In fact the spatial invariance of this problem can almost be ascertained from basic physical symmetry arguments, without the need to write down the underlying PDEs. This model is spatially invariant in the horizontal dimension, but not in the vertical one.

The realization that this problem is spatially invariant in the horizontal dimension significantly reduces the complexity of the control design. To approximate the linearized Navier-Stokes equations in both directions at large Reynolds numbers is prohibitively expensive computationally, and would yield controllers of very large orders. The horizontal spatial invariance of this problem implies that *a)* one only needs to approximate in the vertical dimension, the horizontal dimension should be Fourier transformed, and *b)* optimal controllers have the structure of horizontal spatial convolutions that are spatially localized (see Section V). We note here that the first fact *a)* was observed in [7] by looking at the modal decomposition, and the second fact was recently utilized [42].

B. A General Result on the Spatial Invariance of Optimal Controllers

In this section we divert from our $\mathcal{L}_2$ set up and consider problems for general $\mathcal{L}_p$ induced norms. We have seen that quadratically optimal controllers for spatially invariant systems are themselves spatially invariant. We will ask a similar question for more general $\mathcal{L}_p$ induced norms: given a spatially-invariant generalized plant, can the optimal controller be taken to be spatially invariant? Fortunately, the answer is yes, and this implies a significant reduction in the complexity of the control design problem.

The question above is reminiscent of questions related to time-varying versus time-invariant compensation [43]–[46], where it was shown that for linear time invariant plants and induced norm performance objectives, time-varying controllers offer no advantage over time-invariant ones. We will now prove a similar result. The essence of our technique is similar to that of [45], [46], in which averaging over the time index was used. We will show here that the averaging argument only needs a group structure, and thus one can average over any group, in particular our spatial index sets.

To state the result precisely, consider a set up in terms of the "standard problem" of robust control [38], [47], [48]. This is shown in Fig. 1. All signals w, z, u, y are vector-valued signals indexed over the same group $\mathbb{G}$. The objective in this problem is to find a stabilizing controller that minimizes the $\mathcal{L}_p(\mathbb{G})$-induced norm from from w to z. For reference we note here that the $\mathcal{L}_p(\mathbb{G})$ norm is defined as

$$\|w\|_p := \left(\int_{\mathbb{G}}\int_{\mathbb{R}} |w_x(t)|^p dt dx\right)^{1/p} = \left(\int_{\mathbb{G}} \|w\|_p^p\right)^{1/p}.$$

We note that taking such norms on the regulated variables z allows us to penalize *global* objectives, i.e., to design the controller so that some macro-objective of the overall array system is optimized.

In the usual notation, we will refer to the closed loop system in Fig. 1 by $\mathcal{F}(G,K)$. For any given $p \in (0,\infty]$, the input-output sensitivity of the system is given by the $\mathcal{L}_p$ induced norm

$$\|\mathcal{F}(G,K)\|_{p-i} := \sup_{w\in\mathcal{L}_p} \frac{\|z\|_p}{\|w\|_p}.$$

If the controller is a stabilizing controller the above worst case gain will be finite.

Let *LSI* and *LSV* be the classes of Linear Spatially Invariant and Linear Spatially Varying (not necessarily stable) systems respectively. let us define the following two problems:

$$\gamma_{si} := \inf_{\substack{\text{stabilizing}\\ K\in LSI}} \|\mathcal{F}(G,K)\|_{p-i},$$
$$\gamma_{sv} := \inf_{\substack{\text{stabilizing}\\ K\in LSV}} \|\mathcal{F}(G,K)\|_{p-i},$$

which are the best achievable performances with *LSI* and *LSV* controllers respectively.

Theorem 9: If the generalized plant G is spatially invariant and has at least one spatially-invariant stabilizing controller, then the best achievable performance can be approached with a spatially-invariant controller. More precisely,

$$\gamma_{si} = \gamma_{sv}.$$

Proof: We use the existence of one stabilizing controller to obtain a stable coprime factorization of the plant [4, chap. 8]. Since the controller is spatially invariant, then so are the stable factors. With the YJBK parameterization, our problem then becomes

$$\gamma_{si} = \inf_{\substack{\text{stable}Q\\ Q\in LSI}} \|H_1 - H_2 Q H_3\|$$
$$\gamma_{sv} = \inf_{\substack{\text{stable}Q\\ Q\in LSV}} \|H_1 - H_2 Q H_3\|$$

where the stable systems $\{H_i\}$ are LSI.

We first consider the case where the group $\mathbb{G}$ is finite. Let $\bar{Q}$ (possibly LSV) achieve a given performance level $\bar{\gamma} = \|H_1 - H_2\bar{Q}H_3\|$. Define the averaged system

$$Q_a := \frac{1}{|\mathbb{G}|}\sum_{\sigma\in\mathbb{G}} T_{-\sigma}\bar{Q}T_\sigma,$$

where the "shift" operator T_σ acts by $(T_\sigma\psi)(x) := \psi(x-\sigma)$. Now it is clear that Q_a is spatially invariant, since the group property guarantees for any $x \in \mathbb{G}$

$$T_{-x}Q_aT_x = T_{-x}\left(\sum_{\sigma\in\mathbb{G}} T_{-\sigma}\bar{Q}T_\sigma\right)T_x = \sum_{\sigma\in\mathbb{G}} T_{-\sigma}\bar{Q}T_\sigma$$

(i.e., the sum is re-shuffled). Now we prove that Q_a achieves at least the same performance as $\bar{Q}$. This is standard since

$$\begin{aligned}\|H_1 - H_2Q_aH_3\| &= \left\|H_1 - H_2(\frac{1}{|\mathbb{G}|}\sum_{\sigma\in\mathbb{G}} T_{-\sigma}\bar{Q}T_\sigma)H_3\right\|\\ &= \left\|\frac{1}{|\mathbb{G}|}\sum_{\sigma\in\mathbb{G}} T_{-\sigma}\left(H_1 - H_2\bar{Q}H_3\right)T_\sigma\right\|\\ &\le \frac{1}{|\mathbb{G}|}\sum_{\sigma\in\mathbb{G}} \left\|T_{-\sigma}\left(H_1 - H_2\bar{Q}H_3\right)T_\sigma\right\|\\ &= \|H_1 - H_2\bar{Q}H_3\|,\end{aligned}$$

where we have used the spatial invariance of the $\{H_i\}$'s and of the norm.

In the case of infinite $\mathbb{G}$, we follow the argument of [46] with a slight modification. We take a sequence of subsets

$$M_1 \subset M_2 \subset \cdots \text{ with } \bigcup_n M_n = \mathbb{G},$$

where each M_n has finite Haar measure $|M_n|$ (see [31]), and define

$$Q_n := \frac{1}{|M_n|}\int_{M_n} T_{-\sigma}\bar{Q}T_\sigma d\sigma.$$

This sequence $\{Q_n\}$ then converges weak-$*$ to a spatially-invariant Q_a with the required norm bound [46]. ■

The above result can be seen as a sequel to that in [27], [28] showing existence of stabilizing controllers with the same invariance property as the plant. We have extended this here to *optimal* controllers.

Remark: Although the examples we consider in this paper are over commutative groups, the proof of the above theorem does not require this assumption, and the statement holds for noncommutative groups as well. For example, a MIMO transfer function matrix where the entries are invariant with respect to a permutation subgroup of the input and output indices will have an optimal controller with the same invariance property.

The practical implication of this result is that if the plant is spatially invariant (which is often obvious from physical symmetries), the controllers can be taken to be spatial convolutions. We note that this also applies to systems with an *uneven* distribution of sensors and actuators. In such systems, if there is a large number of sensors and actuators, they are typically distributed in a regular lattice structure. Since any regular lattice can be generated by shifts of a fundamental cell (see [36] for an illustrative example), such a lattice exhibits a translational invariance with respect to those shifts. Our results then specify what shift invariance optimal controllers can have. Such a structure may be useful even when implementing controllers designed using other methods (e.g., networks of PID controllers).

VII. Conclusion

In this paper we have studied control problems with distributed sensing and actuation over spatial coordinates. We have identified the spatial invariance property as a fundamental tool in reducing the complexity of optimal control design with regard to global objectives. In the special case of quadratic performance measures (LQR, $\mathcal{H}_2$, $\mathcal{H}_\infty$), we provided exact solutions to infinite dimensional control design problems in terms of parametrized families of finite dimensional ones. We have also discovered that these optimal controllers are inherently distributed and spatially localized. Finally, we proved a general principle that any spatial invariance of the plant is inherited by an optimal controller under a variety of performance criteria.

The spatial invariance assumption involves a certain idealization with respect to practical control problems which have bounded spatial domains. Also, the consideration of distributed actuation over a continuous spatial variable will in practice be implemented with some level of discretization. These idealizations are, however, completely analogous to those usually performed over the time domain: a long but finite time horizon is typically treated by infinite horizon techniques, and fast temporally sampled systems are often approximated by continuous time systems.

In terms of spatial localization, the results of Section V-B on exponential decay rates of the optimal convolution kernel provide a qualitative assessment. Although these results are for unbounded domains, examples involving bounded domains appear to have this inherent localization property as well [14], [50]. The next question would be to attempt to influence the degree of localization in the design. One possible line of attack would be the use of spatio-temporal weighting functions (e.g., in $\mathcal{H}_2$ or $\mathcal{H}_\infty$ optimal control); this remains open for further research. Another recently explored strategy [51] involves relaxations to the LMI versions of these problems.

Appendix A
Completing the Proof of Theorem 6

The difficulty in establishing the existence of a solution $P(\sigma)$ for $\sigma \neq j\lambda$ is that the corresponding ARE is not standard, in that the coefficients $-B(\sigma)B^\sim(\sigma)$, $Q(\sigma)$ are not Hermitian. However the Hamiltonian approach to this problem, as described in [52, Ths. 12.1 and 12.2], can be mimicked under our assumptions, to show existence of a solution satisfying $P^\sim(\sigma) = P(\sigma)$. We now sketch this argument, focusing for brevity on the steps which are not straightforward, where our assumptions must be invoked. We first find bases for the stable eigenspaces of $H(\sigma)$ and $H^\sim(\sigma)$ such that

$$H(\sigma)\begin{bmatrix} X_1(\sigma) \\ X_2(\sigma)\end{bmatrix} = \begin{bmatrix} X_1(\sigma) \\ X_2(\sigma)\end{bmatrix} H_-(\sigma)$$
$$\begin{bmatrix} X_1^T(-\sigma) & X_2^T(-\sigma)\end{bmatrix} H^T(-\sigma) = H_-^T(-\sigma) \times \begin{bmatrix} X_1^T(-\sigma) X_2^T(-\sigma)\end{bmatrix} \tag{49}$$

where $H_-(\sigma)$ and $H_-^T(-\sigma)$ are stable matrices. Note that the matrices $X_i(\sigma)$ and $X_i(-\sigma)$ are not necessarily related, but we now derive some relations. Proceeding as in [52, Th. 12.1] and using the stability of $H_-(\sigma)$ and $H_-^T(-\sigma)$, we conclude that

$$X_1^T(-\sigma)X_2(\sigma) = X_2^T(-\sigma)X_1(\sigma) \tag{50}$$

which also implies that $X_1(\sigma)$ is invertible if and only if $X_1(-\sigma)$ is invertible. To see that, assume $X_1(\sigma)$ is invertible and $X_1^T(-\sigma)$ is not. This means $\exists z \in \mathbb{C}^n$ such that $z^T X_1^T(-\sigma) = 0$, so (50) implies that $z^T X_2^T(-\sigma)X_1(\sigma) = 0$. Since $X_1(\sigma)$ is invertible, this means $z^T X_2^T(-\sigma) = 0$, contradicting the fact that $[X_1^T(-\sigma) \quad X_2^T(-\sigma)]$ forms a basis.

These facts imply that when $X_1(\sigma)$ is invertible

$$P(\sigma) := X_2(\sigma)X_1^{-1}(\sigma) = \left(X_2(-\sigma)X_1^{-1}(-\sigma)\right)^T$$

is well defined and has the property $P^\sim(\sigma) = P(\sigma)$.

It remains to show that $X_1(\sigma)$ is invertible. Suppose it is not, then we first have

Claim: Either $\ker(X_1(\sigma))$ is $H_-(\sigma)$ invariant, or $\ker(X_1(-\sigma))$ is $H_-(-\sigma)$ invariant.

Proof: Suppose not, then $\exists\ y, z \in \mathbb{C}^n$ such that $z^T X_1^T(-\sigma) = 0$ & $z^T H_-^T(-\sigma)X_1^T(-\sigma) \neq 0$, and $X_1(\sigma)y = 0$ & $X_1(\sigma)H_-(\sigma)y \neq 0$. Now, the first component of (49) is

$$A(\sigma)X_1(\sigma) - B(\sigma)B^\sim(\sigma)X_2(\sigma) = X_1(\sigma)H_-(\sigma) \tag{51}$$

which then gives

$$\begin{aligned} &z^T X_2^T(-\sigma)X_1(\sigma)H_-(\sigma)y \\ &\quad = -z^T \times X_2^T(-\sigma)B(\sigma)B^T(-\sigma)X_2(\sigma)y \\ &\quad = z^T X_1^T(-\sigma)X_2(\sigma)H_-(\sigma)y = 0 \end{aligned}$$

which by our assumptions imply that either $z^T X_2^T(-\sigma)B(\sigma) = 0$ or $B^T(-\sigma)X_2(\sigma)y = 0$ (or both). This last statement, together with (51) and its "dual" imply

$$\begin{aligned} B^T(-\sigma)X_2(\sigma)y &= 0 \Rightarrow X_1(\sigma)H_-(\sigma)y = 0 \\ z^T X_2^T(-\sigma)B(\sigma) &= 0 \Rightarrow z^T H_-^T(-\sigma)X_1^T(-\sigma) = 0. \end{aligned}$$

Since at least one of the above is true, we have a contradiction, and the claim is proved. ■

We will assume the first clause of the claim and show that the stabilizability assumption is then violated, a symmetrical argument can be made if the second clause is true.

Since $\ker(X_1(\sigma))$ is $H_-(\sigma)$ invariant, $\exists\ v \in \ker(X_1(\sigma))$ such that $H_-(\sigma)v = \lambda v$, with $\mathrm{Re}(\lambda) < 0$. Multiplying (49) by v, we obtain

$$\begin{bmatrix} -B(\sigma)B^\sim(\sigma) \\ -A^T(-\sigma)\end{bmatrix} X_2(\sigma)v = \begin{bmatrix} 0 \\ \lambda X_2(\sigma)v\end{bmatrix}$$
$$\Downarrow$$
$$\begin{bmatrix} -B(\sigma)B^\sim(\sigma) \\ -A^T(-\sigma) - \lambda I\end{bmatrix} X_2(\sigma)v = 0$$

which violates the stabilizability assumption (at $-\sigma$) since $X_2(\sigma)v$ is a nonzero vector (because X_1, X_2 form a basis), and $\mathrm{Re}(-\lambda) > 0$.

We have thus established the existence of a solution with the property that $P^{\sim}(\sigma) = P(\sigma)$. This solution gives a stable $A(\sigma) - B(\sigma)B^{\sim}(\sigma)P(\sigma)$ by a standard argument as in [52, Th. 12.1].

APPENDIX B
MATRIX AREs AND ALGEBRAIC FUNCTIONS

In applying the results of Theorem 5 to exponential decay of optimal Riccati operators, we invoked the *algebraic* property of the solution $p(\sigma)$ in the scalar case. We want to extend this to the matrix case; here the ARE is in effect a system of $m = n^2$ coupled equations in the n^2 entries of P and the variable σ; under Assumption 2 we can eliminate denominators and write the equivalent polynomial equations

$$g_k(z_1, \ldots, z_m, \sigma) = 0, \qquad k = 1 \ldots m \tag{52}$$

where $z_1, \ldots, z_m$ are the entries p_{ij} of P.

Let $\check{z}_1(\sigma), \ldots, \check{z}_m(\sigma)$ be the entries of the stabilizing solution of the ARE; we claim that each of these is a scalar algebraic function of σ to which we can apply Proposition 8. We outline a proof based on some tools from *algebraic geometry*, taken from [53].

- The set $V \subset \mathbb{C}^{m+1}$ of solutions to equations such as (52) is called an *affine variety*. Algebraic geometry studies such sets by means of their interplay with *ideals* in the ring of polynomials $\mathbb{C}[z_1, \ldots, z_m, \sigma]$. An ideal generated by polynomials $g_1, \ldots, g_s$ is the set

  $$I = \{h_1 g_1 + \cdots + h_s g_s, \quad h_i \in \mathbb{C}[z_1, \ldots, z_m, \sigma]\}$$

  the Hilbert basis theorem states all ideals are of this form. If we consider polynomials g_i defining a variety V, then $V = \mathbf{V}(I)$ denotes the fact that V is the common set of roots of polynomials in I. All varieties have an associated ideal. Also, a variety can be decomposed into the union of a finite number of "irreducible" varieties, which have a well-defined *dimension*: at all but a set of singular points, this coincides with the dimension of the tangent subspace.
- The intersection of an ideal I with the ring $\mathbb{C}[z_{k+1}, \ldots, z_m, \sigma]$ is called the elimination ideal I_k; it contains equations that can be algebraically derived from those in I and which eliminate the first k variables. If we now look at the set $V(I_k)$ of roots $(z_{k+1}, \ldots, z_m, \sigma)$ to the polynomials in the elimination ideal this must contain the projection $\pi_k(V)$ of V onto the coordinates $z_{k+1}, \ldots, z_m, \sigma$; indeed $V(I_k)$ is the "closure" of this set in an algebraic sense.

We now state the following result.

Proposition 10: Consider the polynomial (52). Suppose that for σ in an open set $\mathcal{S} \subset \mathbb{C}$, we have a continuous family of solutions $\check{z}_1(\sigma), \ldots, \check{z}_m(\sigma)$, such that the Jacobian matrix

$$\frac{\partial g}{\partial z}$$

is nonsingular at every point $(\check{z}_1(\sigma), \ldots, \check{z}_m(\sigma), \sigma)$. Then there exists a polynomial

$$h(z, \sigma) = \gamma_d(\sigma) z^d + \cdots + \gamma_1(\sigma) z + \gamma_0(\sigma) \tag{53}$$

where $d > 0$, $\gamma_d(\sigma) \not\equiv 0$, satisfying $h(\check{z}_m(\sigma), \sigma) \equiv 0$.

Remark: We chose z_m for concreteness, but an analogous argument can be used for the other variables.

Proof: Let $V \subset C^{m+1}$ be the variety defined by the (52). Invoking [53, Th. 9, p. 462], we find that all points $(\check{z}_1(\sigma), \ldots, \check{z}_m(\sigma), \sigma)$ are nonsingular and belong to a unique irreducible component of V of dimension one. Let V^0 be this irreducible component (which must be the same for all points by continuity), and I^0 the corresponding prime ideal.

Now, we consider the elimination ideal $I^0_{m-1} = I^0 \cap \mathbb{C}[z_m, \sigma]$; since V^0 has dimension one, we see [53, Cor 4, p. 449] that I^0_{m-1} must be nontrivial, and thus we find a nontrivial $h(z_m, \sigma)$ that is zero on V^0; this implies $h(\check{z}_m(\sigma), \sigma) \equiv 0$. This polynomial must depend explicitly on z_m, otherwise, we would have derived a constraint on σ, which we assumed is free to vary in an open set. ∎

We can now apply this result to the $m = n^2$ polynomial equations obtained from the ARE (41)

$$\begin{aligned} 0 =& F(\sigma, P) \\ :=& \phi(\sigma) \left(A^{\sim}(\sigma) P + P A(\sigma) - P B(\sigma) B^{\sim}(\sigma) P + Q(\sigma) \right). \end{aligned}$$

Here $\phi(\sigma)$ is the common denominator of $A(\sigma)$, $B(\sigma)$, $Q(\sigma)$; note that by Assumption 2 (i), $\phi(\sigma)$ is nonzero on $\mathcal{S} = \{\sigma \in \mathbb{C} : |\mathrm{Re}(\sigma)| < \alpha\}$.

We know from Appendix A that for each $\sigma \in \mathcal{S}$, there is a solution $\check{P}(\sigma)$ which is stabilizing, i.e., $(A(\sigma) - B(\sigma)B^{\sim}(\sigma)\check{P}(\sigma))$ is a stable matrix. By symmetry of the region the same happens with $(A^{\sim}(\sigma) - \check{P}(\sigma)B(\sigma)B^{\sim}(\sigma))$. We must show the Jacobian of these equations is nonzero on $\check{P}(\sigma)$; we can in fact express the partial derivatives $\partial F / \partial p_{ij}$ in matrix form (in vector form they would correspond to one column of the Jacobian), as

$$\frac{\partial F}{\partial p_{ij}} = \phi(\sigma) \left((A^{\sim} - P B B^{\sim}) E_{ij} + E_{ij} (A - B B^{\sim} P) \right)$$

where E_{ij} is the canonical matrix basis. If the Jacobian matrix is singular at $(\sigma, \check{P}(\sigma))$, we can find scalars m_{ij}, not all zero, such that

$$\sum_{i,j} m_{ij} \frac{\partial F}{\partial p_{ij}} (\sigma, \check{P}) = 0$$

in matrix form, this combination takes the form

$$\phi(\sigma) \left((A^{\sim} - \check{P} B B^{\sim}) M + M (A - B B^{\sim} \check{P}) \right) = 0$$

where M is the matrix of the m_{ij}; but now the stability of the matrices $(A - BB^{\sim}\check{P})$ and $(A^{\sim} - \check{P}BB^{\sim})$ implies that $M = 0$ invoking Sylvester's theorem. So, we have a contradiction and the Jacobian is nonsingular.

Now, Proposition 10 implies that each $\check{p}_{ij}$ is algebraic, and we can then obtain a growth bound by applying the scalar result in Proposition 8.

Remark: The nonsingularity of the Jacobian has an additional implication; invoking the implicit function theorem, we see that the stabilizing solution $\check{P}(\sigma)$ will be differentiable as a function of the real and imaginary parts of σ; this fact is used in the proof of Theorem 6.

ACKNOWLEDGMENT

The first author would like to thank P. Kokotovic and M. Dahleh for providing a supportive research environment while visiting at the Center for Control Engineering and Computations at the University of California, Santa Barbara, during 1997–98. The authors would like to thank the anonymous referees and H. Ozbay for identifying technical issues in need of improvement, and P. Parrilo for helpful references.

REFERENCES

[1] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*. New York: Springer-Verlag, 1971.

[2] S. P. Banks, *State-Space and Frequency-Domain Methods in the Control of Distributed Parameter Systems*. London, U.K.: Peter Peregrinus, 1983.

[3] R. F. Curtain and A. J. Pritchard, "Infinite dimensional linear systems theory," *Lecture Notes Control Inform. Sci.*, vol. 8, 1978.

[4] A. Bensousan, G. P. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and Control of Infinite Dimensional Systems*. Boston, MA: Birkhauser, 1992 & 1993, vol. I & II.

[5] C. Liu, T. Tsao, Y. Tai, J. Leu, C. Ho, W. Tang, and D. Miu, "Out of plane permanent magnetic actuators for delta wing motion control," in *Proc. 1995 IEEE Micro Electro Mechanical Systems Workshop*, Amsterdam, The Netherlands, 1995.

[6] C. Ho and Y. Tai, "REVIEW: MEMS and its applications for flow control," *ASME J. Fluid Eng.*, vol. 118, Sept. 1996.

[7] S. S. Joshi, J. L. Speyer, and J. Kim, "A system theory approach to the feedback stabilization of infinitesimal and finite-amplitude distrubances in plane Poiseuille flow," *J. Fluid Mech.*, 1997.

[8] T. R. Bewley and S. Liu, "Optimal and robust control and estimation of linear paths to transition," *J. Fluid Mech.*, 1998.

[9] H. T. Banks, R. S. Smith, and Y. Wang, *Smart Material Structures: Modeling, Estimation and Control*. New York: Wiley, 1996.

[10] R. E. Skelton and C. Sultan, "Controllable tensegrity: A new class of smart structures," in *Proc. SPIE 4th Annual Int. Symp. Smart Structures Materials*, San Diego, CA, Mar. 1997.

[11] A. V. Balakrishnan, "Aeroelastic control with self-straining actuators: Continuum models," in *Proc. SPIE Fifth Annual Int. Symp. Smart Structures Materials*, San Diego, CA, Mar. 1998.

[12] S. M. Melzer and B. C. Kuo, "Optimal regulation of systems described by a countably infinite number of objects," *Automatica*, vol. 7, pp. 359–366, 1971.

[13] K. C. Chu, "Decentralized control of high-speed vehicular strings," *Trans. Sci.*, pp. 361–384, Nov. 1974.

[14] B. Shu, "Robust Longitudinal Control of Vehicle Platoons on Intelligent Highways," M.S. thesis, Dept. ECE, Univ. Illinois, Urbana-Champaign, 1996.

[15] E. M. Heaven, I. M. Jonsson, T. M. Kean, M. A. Manness, and R. N. Vyse, "Recent advances in cross machine profile control," *IEEE Control Syst. Mag.*, vol. 14, no. 5, Oct. 1994.

[16] D. Laughlin, M. Morari, and R. D. Braatz, "Robust performance of cross-directional basis-weight control in paper machines," *Automatica*, vol. 29, pp. 1395–1410, 1993.

[17] R. D. Braatz and J. G. VanAntwerp, "Robust cross-directional control of large scale paper machines," in *Proc. 1996 IEEE Int. Conf. Control Applications*, 1996.

[18] E. W. Kamen, "Stabilization of linear spatially-distributed continuous-time and discrete-time systems," in *Multidimensional Systems Theory*, N. K. Bose, Ed. Norwell, MA: Kluwer, 1985.

[19] M. L. El-Sayed and P. S. Krishnaprasad, "Homogeneous interconnected systems: An example," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 894–901, Apr. 1981.

[20] R. W. Brockett and J. L. Willems, "Discretized partial differential equations: Examples of control systems defined on modules," *Automatica*, vol. 10, pp. 507–515, 1974.

[21] P. P. Khargonekar, K. Poolla, and A. Tannenbaum, "Robust control of linear time-invariant plants using periodic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-30, Nov. 1985.

[22] J. A. Burns and D. Rubio, "A distributed parameter control approach to sensor location for optimal feedback control of thermal processes," in *Proc. 36th IEEE Conf. Decision Control*, Dec. 1997.

[23] J. A. Burns and B. B. King, "Optimal sensor location for robust control of distributed parameter systems," in *Proc. 33rd IEEE Conf. Decision Control*, Lake Buena Park, FL, Dec. 1994.

[24] W. L. Green and E. W. Kamen, "Stability of linear systems over a commutative normed algebra with applications to spatially-distributed and parameter-dependent systems," *SIAM J. Control Optimiz.*, vol. 23, pp. 1–18, 1985.

[25] E. W. Kamen and P. P. Khargonekar, "On the control of linear systems whose coefficients are functions of parameters," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 25–33, Jan. 1984.

[26] P. P. Khargonekar and E. Sontag, "On the relation between stable matrix fraction factorizations and regulable realizations of linear systems over rings," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 627–638, Mar. 1982.

[27] F. Fagnani and J. C. Willems, "Representations of symmetric linear dynamical systems," *SIAM J. Control Optimiz.*, vol. 31, no. 5, pp. 1267–1293, 1993.

[28] ——, "Interconnections and symmetries of linear differential systems," *SIAM J. Control Optimiz.*, 1994.

[29] R. D'Andrea, "A linear matrix inequality approach to decentralized control of distributed parameter systems," in *Proc. Amer. Control Conf.*, June 1998.

[30] G. Dullerud, R. D'Andrea, and S. Lall, "Control of spatially varying distributed systems," in *Proc. IEEE Control Decision Conf.*, Dec. 1998.

[31] W. Rudin, *Fourier Analysis on Groups*. New York: Interscience-Wiley, 1962.

[32] R. Walter, *Functional Analysis*. New York: McGraw-Hill, 1991.

[33] R. F. Curtain and H. J. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*. New York: Springer-Verlag, 1995.

[34] B. van Keulen, *$\mathcal{H}_\infty$-Control for Distributed Parameter Systems*. Boston, MA: Birkhauser, 1993.

[35] J. C. Doyle, K. Glover, P. Khargonekar, and B. A. Francis, "State-space solutions to standard $\mathcal{H}^2$ and $\mathcal{H}^\infty$ problems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 831–847, Aug. 1989.

[36] B. Bamieh, F. Paganini, and M. A. Dahleh. (2001) Distributed Control of Spatially Invariant Systems. [Online]. Available: www.engineering.ucsb.edu/~bamieh/papers/ccec01-0720.ps.

[37] L. Hormander, *The Analysis of Partial Differential Operators I*, 2nd ed. Berlin, Germany: Springer-Verlag, 1990.

[38] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[39] J. H. Davenport, Y. Siret, and E. Tournier, *Computer Algebra, Systems and Algorithms for Algebraic Computation*. New York: Academic, 1988.

[40] B. Mishra, *Algorithmic Algebra*. New York: Springer-Verlag, 1993.

[41] A. Packard and J. C. Doyle, "The complex structured singular value," *Automatica*, vol. 29, no. 1, pp. 71–109, 1993.

[42] H. Markus and B. Thomas, "Spatially localized convolution kernels for feedback control of transitional flows," in *Proc. 39th IEEE Conf. Decision Control*, Dec. 2000.

[43] A. Feintuch and B. A. Francis, "Uniformly optimal control of linear time-varying systems," *SIAM J. Cont. Optimiz.*, 1985.

[44] P. Khargonekar and K. Poolla, "Uniformly optimal control of linear time-invariant plants: Nonlinear time-varying controllers," *Syst. Control Lett.*, vol. 6, pp. 303–308, Jan. 1986.

[45] J. Shamma and M. Dahleh, "Time-varying versus time-invariant compensation for rejection of persistent bounded disturbances and robust stabilization," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 838–748, July 1991.

[46] H. Chapellat and M. Dahleh, "Analysis of time-varying control strategies for optimal disturbance rejection and robustness," *IEEE Trans. Automat. Contr.*, vol. 37, Nov. 1992.

[47] M. A. Dahleh and I. J. Diaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Upper Saddle River, NJ: Prentice-Hall, 1995.

[48] M. Green and D. Limebeer, *Linear Robust Control*. New York: Prentice-Hall, 1995.

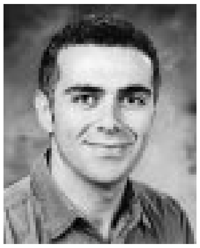
[49] M. Vidyasagar, *Control Systems Synthesis: A Factorization Approach*. Cambridge, MA: MIT Press, 1995.

[50] G. Hagen, I. Mezić, B. Bamieh, and Z. Kaixia, "Control of axial compressors via air injection," in *Proc. Amer. Control Conf.*, San Diego, CA, June 1999.

[51] G. Ayres and F. Paganini, "Convex method for decentralized control design in spatially invariant systems," in *Proc. 39th IEEE Conf. Control*, Dec. 2000.

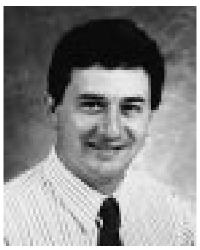
[52] K. Zhou, *Essentials of Robust Control*. Upper Saddle River, NJ: Prentice-Hall, 1998.

[53] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. New York: Springer-Verlag, 1992.



Bassam Bamieh (S'94–M'90) received the electrical engineering and physics degrees from Valparaiso University, Valparaiso, IN, and the M.Sc. and Ph.D. degrees from Rice University, Houston, TX, in 1993, 1996, and 1992, respectively. Between 1991–1998, he was with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. He joined the Mechanical Engineering department at the University of California at Santa Barbara in 1998, where he is now an Associate Professor. His current research interests are in distributed systems, shear flow turbulence modeling and control, Atomic Force Microscopy, multimicro-cantilevers modeling and control, and optical actuation via optical tweezers.

Dr. Bamieh is a past receipient of the AACC Hugo Schuck Best Paper Award and a National Science Foundation CAREER Award.



Fernando Paganini (S'90–M'90) received the electrical engineering and mathematics degrees from the Universidad de la Republica, Montevideo, Uruguay, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1990, 1992, and 1996, respectively.

From 1996 to 1997, he was a Postdoctoral Associate at the Massachusetts Institute of Technology, Cambridge. Since 1997, he has been with the Electrical Engineering Department at the University of California, Los Angeles, where he is currently Associate Professor. His research interests are robust control, distributed control, and networks.

Dr. Paganini received the American Automatic Control Council O. Hugo Schuck award in 1995, the Wilts and Clauser Prizes for outstanding Ph.D. dissertation at the California Institute of Technology in 1996, the 1999 National Science Foundation CAREER Award, and the 1999 Packard Fellowship.

Munther A. Dahleh (S'84–M'97–SM'97–F'01) received the B.S. degree from Texas A & M University, College Station, and the Ph.D. degree from Rice University, Houston, TX, both in electrical engineering, in 1983 and 1987, respectively.

Since 1987, he has been with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, where he is now a Full Professor. He was a Visiting Professor at the Department of Electrical Engineering, California Institute of Technology, Pasadena, in the spring 1993. He has held consulting positions with several companies in the United States and abroad. He is the coauthor of *Control of Uncertain Systems: A Linear Programming Approach* (Upper Saddle River, NJ: Prentice-Hall, 1995), and *Computational Methods for Controller Design* (NY: Springer-Verlag, 1998). His interests include robust control and identification, the development of computational methods for linear and nonlinear controller design, and applications of feedback control in several disciplines including material manufacturing and modeling of biological systems.

Dr. Dahleh was the recipient of the Ralph Budd Award for the best thesis at Rice University in 1987, the George Axelby Outstanding Paper Award (paper coauthored with J. B. Pearson in 1987), a National Science Foundation Presidential Young Investigator Award in 1991, the Finmeccanica Career Development Chair in 1992, the Donald P. Eckman Award from the American Control Council in 1993, and the Graduate Students Council Teaching Award in 1995. He was a Plenary Speaker at the 1994 American Control Conference, and is currently serving as an Associate Editor for IEEE TRANSACTIONS ON AUTOMATIC CONTROL.