Optimization With Affine Homogeneous Quadratic Integral Inequality Constraints

Giovanni Fantuzzi, Andrew Wynn, Paul J. Goulart, *Member, IEEE*, and Antonis Papachristodoulou

***Abstract*—We introduce a new technique to optimize a linear cost function subject to an affine homogeneous quadratic integral inequality, i.e., the requirement that a homogeneous quadratic integral functional affine in the optimization variables is nonnegative over a space of functions defined by homogeneous boundary conditions. Such problems arise in control and stability or input-to-state/output analysis of systems governed by partial differential equations (PDEs), particularly fluid dynamical systems. We derive outer approximations for the feasible set of a homogeneous quadratic integral inequality in terms of linear matrix inequalities (LMIs), and show that a convergent sequence of lower bounds for the optimal cost can be computed with a sequence of semidefinite programs (SDPs). We also obtain inner approximations in terms of LMIs and sum-of-squares constraints, so upper bounds for the optimal cost and strictly feasible points for the integral inequality can be computed with SDPs. We present QUINOPT, an open-source add-on to YALMIP to aid the formulation and solution of our SDPs, and demonstrate our techniques on problems arising from the stability analysis of PDEs.**

***Index Terms*—Integral inequalities, partial differential equations (PDEs), semidefinite programming, sum-of-squares optimization.**

I. Introduction

ANALYSIS and control of systems governed by partial differential equations (PDEs) are fundamental problems in physics and engineering, but are challenging because the system state is a (vector-valued) function $\boldsymbol{w}$ of both the time t and the spatial position vector $\boldsymbol{x}$, and as such it belongs to an infinite-dimensional function space (e.g., a Sobolev space).

In an effort to reduce the conservativeness introduced by finite-dimensional approximations, recent years have seen the development of analytical techniques that consider directly the infinite-dimensional PDEs, and lead to consideration of integral inequalities. For example, the stability of an equilibrium of a PDE system in a domain Ω, or of a control policy designed to stabilize it, can be established by constructing a positive integral Lyapunov functional $\mathcal{V}(t) = \mathcal{V}\{\boldsymbol{w}(t,\cdot)\} = \int_\Omega V[\boldsymbol{w}(t,\boldsymbol{x})]\mathrm{d}\boldsymbol{x}$ whose time derivative (also an integral quantity) is nonpositive [1]–[3]. Other input-to-state/output properties, such as passivity, reachability, and input-to-state stability, can be studied in a similar way using dissipation inequalities for integral functionals of the state variable [4], [5]. Finally, the computational cost of designing optimal control policies for systems with complex dynamics, such as turbulent flows, may be reduced by requiring the control law to minimize an upper bound on the objective function rather than the objective itself [6]–[9], and in the case of PDEs such upper bounds can be found by solving suitable integral inequalities [10]–[15].

When the underlying PDE system is *autonomous*, the integral inequalities obtained in all aforementioned applications depend on time only through the state $\boldsymbol{w}(t,\boldsymbol{x})$, and since they are imposed *pointwise in time*, the time dependence of $\boldsymbol{w}$ can be dropped. Checking a certain integral inequality for given system parameters, or alternatively optimizing the system parameters while satisfying an integral inequality, then requires solving optimization problems of the form

$$\min_{\boldsymbol{\gamma}} \quad \boldsymbol{c}^T\boldsymbol{\gamma}$$
$$\text{s.t.} \quad \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \int_\Omega F_{\boldsymbol{\gamma}}(\boldsymbol{x}, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w})\mathrm{d}^n\boldsymbol{x} \geq 0, \quad \boldsymbol{w} \in H \qquad (1)$$

where H is a suitable function space, e.g., the space of all $\boldsymbol{k}$-times differentiable functions from $\Omega \subseteq \mathbb{R}^n$ (typically $n = 3$ for physical systems) to $\mathbb{R}^q$ that satisfy a given set of boundary conditions (BCs). The optimization variable $\boldsymbol{\gamma} \in \mathbb{R}^s$ represents a vector of tunable system parameters, $\boldsymbol{c} \in \mathbb{R}^s$ is the cost vector, $F_{\boldsymbol{\gamma}}(\cdot,\cdot)$ is a function that depends parametrically on $\boldsymbol{\gamma}$, and $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w} = [w_1, \partial_{x_1} w_1, \partial_{x_2} w_1, \ldots, \partial_{x_n}^{k_1} w_1, \ldots, \partial_{x_n}^{k_q} w_q]^T$ lists all partial derivatives of the components of $\boldsymbol{w}$ up to the order specified by the multi-index $\boldsymbol{k} = [k_1, \ldots, k_q]$.

When the dependence on $\boldsymbol{\gamma}$ is at least affine and strong duality holds, problem (1) could be solved (in principle) by first computing the minimizer $\boldsymbol{w}^\star$ of $\mathcal{F}_{\boldsymbol{\gamma}}$ as a function of $\boldsymbol{\gamma}$ using the calculus of variations [16], [17], and then minimizing the augmented Lagrangian $L(\boldsymbol{\gamma}) = \boldsymbol{c}^T\boldsymbol{\gamma} - \lambda\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}^\star\}$, where the Lagrange multiplier $\lambda \geq 0$ is chosen to enforce the integral inequality. This strategy has been successfully applied to some problems in fluid dynamics (see, e.g., [18]–[20]), but it requires careful, problem-dependent computations. Alternatively, when the integrand $F_{\boldsymbol{\gamma}}(\cdot,\cdot)$ is linear with respect to $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$ and polynomial in $\boldsymbol{x}$, (1) can be transformed into a semidefinite program (SDP) using integration by parts and moment relaxation tech-

Manuscript received July 20, 2016; revised February 15, 2017; accepted April 26, 2017. Date of publication June 26, 2017; date of current version December 1, 2017. This work was supported in part by Engineering and Physical Sciences Research Council under Grant EP/J010537/1. Recommended by Associate Editor C.-Y. Kao. *(Corresponding author: Giovanni Fantuzzi.)*

G. Fantuzzi and A. Wynn are with the Department of Aeronautics, Imperial College London, South Kensington Campus, London SW7 2AZ, U.K. (e-mail: gf910@ic.ac.uk; a.wynn@imperial.ac.uk).

P. J. Goulart and A. Papachristodoulou are with the Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, U.K. (e-mail: paul.goulart@eng.ox.ac.uk; antonis@eng.ox.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2017.2703927

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/

niques [21]. More recently, it has been suggested that (1) can be recast as an SDP even when the integrand is polynomial in $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$ [2], [3], [22], [23]: one relates the derivatives of the components of $\boldsymbol{w}$ using integration by parts and algebraic identities, and then requires that the polynomial integrand $F_{\boldsymbol{\gamma}}(\boldsymbol{x}, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w})$ admits a sum-of-squares (SOS) decomposition over the domain of integration. However, scalability issues usually prevent the solution of problems of practical interest because—as our examples will demonstrate—high-degree SOS relaxations are normally needed to achieve accurate results.

This paper presents a new approach to solving a class of problems of type (1). We consider homogeneous quadratic functionals $\mathcal{F}_{\boldsymbol{\gamma}}$ over a one-dimensional compact domain; in other words, we assume that $\boldsymbol{x} \in \Omega \equiv [a, b] \subset \mathbb{R}$ and that the integrand $F_{\boldsymbol{\gamma}}(\boldsymbol{x}, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w})$ is a homogeneous quadratic polynomial with respect to $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$. Inequalities of this type arise in many fluid or thermal convection systems of practical interest (see, e.g., [1], [11], [13], [14], [24]) and these are the main applications we have in mind. Our techniques, already partially introduced by some of the authors for particular problem instances [25], [26], rely on Legendre series expansions to formulate SDPs with better scaling properties than the SOS method of [3]. Our main contributions are as follows.

1) For the first time, we formulate convergent *outer* approximations of the feasible set of (1) described by linear matrix inequalities (LMIs), so lower bounds for the optimal cost can be computed using SDPs.
2) We extend the method of [25] and [26] to derive SDP-representable *inner* approximations for the feasible set of (1) in the general setting, so upper bounds on the optimal value of can also be obtained with semidefinite programming.
3) We present QUadratic INtegral OPTimization (QUINOPT), an add-on to YALMIP [27], [28] to aid the formulation of the SDP relaxations outlined above, and use it to solve examples that demonstrate the advantages (and some limitations) of our approach compared to the SOS method of [3].

The rest of the paper is organized as follows. Section II introduces the class of optimization problems studied in this work; as a motivating example, we consider the stability analysis of a fluid flow driven by a surface stress [15]. We formulate outer SDP relaxations in Section III, and inner SDP relaxations in Section IV. We remove some simplifying assumptions and further extend our results in Section V. Section VI presents QUINOPT and numerical examples arising from the analysis of PDEs, and we comment on the scalability of our methods in Section VII. Finally, Section VIII offers concluding remarks and perspectives for future developments.

Notation: Vectors and matrices are denoted by boldface characters; in particular, $\boldsymbol{0}$ denotes the zero vector/matrix. The usual Euclidean and ℓ^1 norms of $\boldsymbol{v} \in \mathbb{R}^n$ are $\|\boldsymbol{v}\| = (\sum_{i=1}^n |v_i|^2)^{1/2}$ and $\|\boldsymbol{v}\|_1 = \sum_{i=1}^n |v_i|$, respectively. Given a matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times m}$, the Frobenius norm is defined as $\|\boldsymbol{Q}\|_F = (\sum_{i=1}^n \sum_{j=1}^m |Q_{ij}|^2)^{1/2}$. The range and null space of $\boldsymbol{Q}$ are denoted by $\mathcal{R}(\boldsymbol{Q})$ and $\mathcal{N}(\boldsymbol{Q})$, respectively. We denote the space of $n \times n$ symmetric matrices by $\mathbb{S}^n$, and indicate that $\boldsymbol{Q} \in \mathbb{S}^n$ is positive semidefinite with the notation $\boldsymbol{Q} \succeq 0$.

For a compact interval $[a, b] \subset \mathbb{R}$ and a positive integer q, $C^m([a, b], \mathbb{R}^q)$ is the space of m-times continuously differentiable functions with domain $[a, b]$ and values in $\mathbb{R}^q$; we also write $C^m([a, b])$ for $C^m([a, b], \mathbb{R})$. Given $u \in C^m([a, b])$, $\|u\|_2$ and $\|u\|_\infty$ denote the usual $L^2(a, b)$ and $L^\infty(a, b)$ norms

$$\|u\|_2 = \left[\int_a^b |u(x)|^2 \,\mathrm{d}x\right]^{1/2}, \qquad \|u\|_\infty = \sup_{x\in[a,b]} |u(x)|.$$

The set of nonnegative integers is denoted by $\mathbb{N}$, and $\mathbb{N}^q$ is the set of multi-indices of the form $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_q]$. The length of the multi-index $\boldsymbol{\alpha} \in \mathbb{N}^q$ is $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_q$. Given $\boldsymbol{w} \in C^m([a, b], \mathbb{R}^q)$ and $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^q$ with $\alpha_i \le \beta_i \le m$ for all $i \in \{1, \ldots, q\}$, we define $\boldsymbol{\beta} - \boldsymbol{\alpha} = [\beta_1 - \alpha_1, \ldots, \beta_q - \alpha_q] \in \mathbb{N}^q$ and we list all multi-index derivatives of order between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in the vector

$$\mathcal{D}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w} := \big[\partial^{\alpha_1} u_1, \ldots, \partial^{\beta_1} u_1, \partial^{\alpha_2} u_2, \ldots, \partial^{\beta_2} u_2, \ldots, \partial^{\beta_q} u_q\big]^T \in \mathbb{R}^{q+|\boldsymbol{\beta}-\boldsymbol{\alpha}|}. \quad (2)$$

We also collect all boundary values of such derivatives in the vector

$$\mathcal{B}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w} := \begin{bmatrix} \mathcal{D}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w}|_a \\ \mathcal{D}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w}|_b \end{bmatrix} \in \mathbb{R}^{2(q+|\boldsymbol{\beta}-\boldsymbol{\alpha}|)}. \quad (3)$$

To simplify the notation, when $\boldsymbol{\alpha} = \boldsymbol{0}$ we will write $\mathcal{D}^{\boldsymbol{\beta}}\boldsymbol{w}$ and $\mathcal{B}^{\boldsymbol{\beta}}\boldsymbol{w}$ instead of $\mathcal{B}^{[\boldsymbol{0},\boldsymbol{\beta}]}\boldsymbol{w}$ and $\mathcal{D}^{[\boldsymbol{0},\boldsymbol{\beta}]}\boldsymbol{w}$.

Finally, given two scalar functions f, g of a scalar variable N, we write $f \sim g$ to indicate that f and g are asymptotically equivalent up to multiplication by a positive constant, that is, $\lim_{N\to\infty} f/g = c$ for some positive constant c.

II. Optimization With Affine Homogeneous Quadratic Integral Inequalities

Let $\boldsymbol{\gamma} \in \mathbb{R}^s$ be a vector of optimization variables, and consider two integers m, q and two multi-indices $\boldsymbol{k} = [k_1, \ldots, k_q]$, $\boldsymbol{l} = [l_1, \ldots, l_q] \in \mathbb{N}^q$, such that

$$1 \le k_i \le m - 1, \qquad i \in \{1, \ldots, q\}, \quad (4a)$$

$$k_i \le l_i \le m, \qquad i \in \{1, \ldots, q\}. \quad (4b)$$

Moreover, let $\boldsymbol{F}_0(x), \ldots, \boldsymbol{F}_s(x) \in \mathbb{S}^{q+|\boldsymbol{k}|}$ be symmetric matrices of polynomials of x of degree at most d_F and define

$$\boldsymbol{F}(x;\boldsymbol{\gamma}) := \boldsymbol{F}_0(x) + \sum_{i=1}^s \gamma_i \boldsymbol{F}_i(x) \quad (5)$$

i.e., $\boldsymbol{F}(x;\boldsymbol{\gamma})$ is a symmetric matrix of polynomials of x of degree at most d_F, the coefficients of which are affine in $\boldsymbol{\gamma}$.

Throughout this paper, we consider linear optimization problems of type (1) subject to *affine homogeneous quadratic integral inequalities*, i.e., problems of the form

$$\min_{\boldsymbol{\gamma}} \quad \boldsymbol{c}^T\boldsymbol{\gamma}$$
$$\text{s.t. } \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \int_{-1}^1 \left(\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\right)^T \boldsymbol{F}(x;\boldsymbol{\gamma})\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\,\mathrm{d}x \ge 0,\ \boldsymbol{w} \in H \quad (6)$$

where $\boldsymbol{c} \in \mathbb{R}^s$ is the cost vector, $\boldsymbol{F}(x;\boldsymbol{\gamma})$ is as in (5), and

$$H := \left\{\boldsymbol{w} \in C^m([-1, 1], \mathbb{R}^q) : \ \boldsymbol{A}\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} = \boldsymbol{0}\right\} \quad (7)$$

is the space of m-times continuously differentiable functions satisfying the p homogeneous BCs defined by the matrix $\boldsymbol{A} \in \mathbb{R}^{p\times 2(q+|\boldsymbol{l}|)}$. There is no loss of generality in fixing the

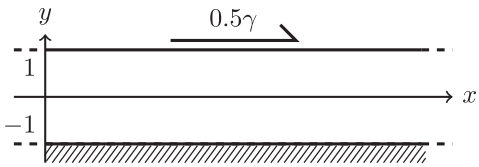


Fig. 1. Sketch of the flow setup in our motivating example. The two-dimensional fluid layer extends to infinity along the x-direction, is bounded at $y = -1$ by a solid boundary and is driven at the surface ($y = 1$) by a shear stress of nondimensional magnitude 0.5γ.

integration domain for the functional $\mathcal{F}_{\boldsymbol{\gamma}}$ to $[-1, 1]$ because any compact interval $[a, b]$ can be mapped to it with a change of integration variable. An affine homogeneous quadratic integral inequality represents a convex constraint on $\boldsymbol{\gamma}$, which makes (6) a convex optimization problem.

Remark 1: For the sake of generality, we allow the space H to be defined by derivatives of higher order than those appearing in $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ (this can always be achieved by adding zero columns to $\boldsymbol{A}$). In the applications we have in mind, i.e., problems arising from the study of autonomous PDEs, this is not uncommon: H encodes the BCs of the solution of a PDE, which might involve all derivatives up to the order of the PDE; $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$, instead, is typically derived from a weak formulation of the PDE, after integrating some terms by parts.

Assumption: To ease the exposition, we only consider two-dimensional (2-D) functions $\boldsymbol{w} = [u, v]^T \in C^m([-1, 1], \mathbb{R}^2)$. We also restrict the attention to the uniform multi-indices $\boldsymbol{k} = [k, k]$ and $\boldsymbol{l} = [l, l]$, where k and l satisfy (4a) and (4b). As will be discussed in Section V, however, all our results hold for the general case.

A. Motivating Example

Consider a 2-D infinite layer of fluid bounded at $y = -1$ by a solid wall and driven at the surface at $y = 1$ by a horizontal shear stress of nondimensional magnitude 0.5γ, as shown in Fig. 1. The flow is governed by the incompressible Navier–Stokes equations, and admits a steady (i.e., time independent) solution in which the flow moves horizontally with velocity $\boldsymbol{w}_0 = (u_0, v_0) = (0.5\gamma y + 0.5\gamma, 0)$; see, for example, [15], [26], [29]. This steady flow is stable when the driving stress is small. The critical value γ_{cr} at which the steady flow is no longer guaranteed to be stable with respect to a sinusoidal perturbation $\boldsymbol{w}(y)e^{i\xi x+\sigma t}$—where $\boldsymbol{w}(y) = [u(y), v(y)]^T$ is the amplitude and ξ is the wave number—is given by the solution of the optimization problem

$$\begin{aligned} &\min \quad -\gamma \\ &\text{s.t.} \quad \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \int_{-1}^{1} \left\{ \frac{16}{\xi^2}[(\partial_y^2 u)^2 + (\partial_y^2 v)^2] + 8[(\partial_y u)^2 \right. \\ &\left. + (\partial_y v)^2] + \xi^2(u^2 + v^2) + \frac{2\gamma}{\xi}(v\partial_y u - u\partial_y v) \right\} \mathrm{d}y \geq 0 \end{aligned} \tag{8}$$

where the integral inequality constraint should hold for all functions $u, v \in C^2([-1, 1])$ satisfying the homogeneous BCs

$$\begin{aligned} u|_{-1} = u|_1 = \partial_y u|_{-1} = \partial_y^2 u|_1 = 0, \\ v|_{-1} = v|_1 = \partial_y v|_{-1} = \partial_y^2 v|_1 = 0. \end{aligned} \tag{9}$$

See [29] and [15] for a detailed discussion. The constraint in (8) can be rewritten in matrix form as in (6) with $\boldsymbol{k} = \boldsymbol{l} = [2, 2]$ and

$$\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w} = \begin{bmatrix} u \\ \partial_y u \\ \partial_y^2 u \\ v \\ \partial_y v \\ \partial_y^2 v \end{bmatrix}, \quad \boldsymbol{F}(x; \boldsymbol{\gamma}) = \begin{bmatrix} \xi^2 & 0 & 0 & 0 & -\frac{\gamma}{\xi} & 0 \\ 0 & 8 & 0 & \frac{\gamma}{\xi} & 0 & 0 \\ 0 & 0 & \frac{16}{\xi^2} & 0 & 0 & 0 \\ 0 & \frac{\gamma}{\xi} & 0 & \xi^2 & 0 & 0 \\ -\frac{\gamma}{\xi} & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{16}{\xi^2} \end{bmatrix}.$$

Note that the matrix $\boldsymbol{F}$ above can be written in the form (5) with $s = 1$. The reader can easily verify that the BCs on u and v can also be rewritten in the matrix form $\boldsymbol{A}\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} = \boldsymbol{0}$ with $\boldsymbol{A} \in \mathbb{R}^{8\times 12}$; we omit the details for brevity. For this problem, it is clear that $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq 0$ for $\gamma = 0$, and that definiteness is lost for sufficiently large γ. However, the interaction of the BCs with this behavior makes the problem interesting and nontrivial to solve. We will compute upper and lower bounds for the optimal γ in (8) in Section VI-A.

III. Outer SDP Relaxations

Our first approach to solve (6) is to derive a sequence of *outer* approximations for its feasible set, defined as

$$T := \{\boldsymbol{\gamma} \in \mathbb{R}^s : \ \forall \boldsymbol{w} \in H, \ \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq 0\}. \tag{10}$$

In other words, we look for a family of sets $\{T_N^{\mathrm{out}}\}_{N\geq 0}$, such that $T \subset T_N^{\mathrm{out}}$. Optimizing the cost function over T_N^{out} then gives a lower bound for the optimal value of (6).

The outer approximation set T_N^{out} can be found by considering a polynomial truncation of $\boldsymbol{w} \in H$ of degree N. In particular, suppose that

$$\boldsymbol{w} = [u, v]^T \in S_N := H \cap (\mathcal{P}_N \times \mathcal{P}_N) \subset H \tag{11}$$

where $\mathcal{P}_N$ is the set of polynomials of degree less than or equal to N on $[-1, 1]$. Note that S_N is nonempty for any degree bound N because H contains the zero polynomial, and it contains nonzero elements if N is large enough to guarantee sufficient degrees of freedom to satisfy the BCs prescribed on H in (7). Finally, $S_N \subset S_{N+1}$ because $\mathcal{P}_N \subset \mathcal{P}_{N+1}$.

Now, let $\hat{u}_0, \ldots, \hat{u}_N$ and $\hat{v}_0, \ldots, \hat{v}_N$ be the coefficients representing the polynomials u and v in any chosen basis for $\mathcal{P}_N$, and define $\boldsymbol{\varphi}_N := [\hat{u}_0, \ldots, \hat{u}_N, \hat{v}_0, \ldots, \hat{v}_N]^T$. Since $\mathcal{F}_{\boldsymbol{\gamma}}$ in (6) is quadratic and the constraints imposed on H are linear, it is clear that there exist a matrix $\boldsymbol{Q}_N(\boldsymbol{\gamma})$, affine in $\boldsymbol{\gamma}$, such that

$$\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} = {\boldsymbol{\varphi}_N}^T \boldsymbol{Q}_N(\boldsymbol{\gamma})\boldsymbol{\varphi}_N$$

and a matrix $\boldsymbol{A}_N$, such that

$$\boldsymbol{w} \in S_N \Leftrightarrow \boldsymbol{A}_N\boldsymbol{\varphi}_N = \boldsymbol{0}.$$

Upon selecting a matrix $\boldsymbol{\Pi}_N$ satisfying $\mathcal{R}(\boldsymbol{\Pi}_N) = \mathcal{N}(\boldsymbol{A}_N)$, it follows that

$$\begin{aligned} T_N^{\mathrm{out}} &:= \{\boldsymbol{\gamma} \in \mathbb{R}^s : \ \forall \boldsymbol{w} \in S_N, \ \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq 0\} \\ &= \left\{\boldsymbol{\gamma} \in \mathbb{R}^s : \ {\boldsymbol{\Pi}_N}^T \boldsymbol{Q}_N(\boldsymbol{\gamma})\boldsymbol{\Pi}_N \succeq 0\right\} \end{aligned} \tag{12}$$

and since $S_N \subset S_{N+1} \subset H$, the feasible set T of (6), defined as in (10), satisfies

$$T \subset T_{N+1}^{\text{out}} \subset T_N^{\text{out}}, \quad N \in \mathbb{N}.$$

This suggests that a sequence of lower bounds on the optimal value of (6) can be found by solving a series of truncated optimization problems.

Theorem 1: Let p^* be the optimal value of (6) and, for each integer N, let p_N^* be the optimal value of the SDP

$$\begin{aligned} \min_{\boldsymbol{\gamma}} \quad & \boldsymbol{c}^T \boldsymbol{\gamma} \\ \text{s.t.} \quad & \boldsymbol{\Pi}_N^T \, \boldsymbol{Q}_N(\boldsymbol{\gamma}) \, \boldsymbol{\Pi}_N \succeq 0. \end{aligned} \tag{13}$$

Then, $\{p_N^\star\}_{N \geq 0}$ is a nondecreasing sequence of lower bounds for $p^\star$. Furthermore, if a minimizer $\gamma^\star$ exists in (6), then $\lim_{N \to \infty} |p_N^\star - p^\star| = 0$.

Proof: See Appendix B-A. ■

Remark 2: Clearly, infeasibility of the SDP (13) for a certain N provides a *certificate of infeasibility* for (6). However, note that the feasibility (resp. unboundedness) of (13) for any finite N does *not* prove that (6) is feasible (resp. unbounded).

It is important to note that Theorem 1 provides no control on the gap $p^\star - p_N^\star$ *as a function of* N. In other words, an arbitrarily large N might be required for a given level of approximation accuracy. Consequently, the rest of this work will focus on proving checkable conditions upon which upper bounds can be placed on p^*.

IV. Inner SDP Relaxations

Upper bounds on the optimal value of (6) that complement the lower bounds from Theorem 1 can be found by optimizing the cost function over an inner approximation T_N^{in} of the true feasible set. Such an inner approximation can be constructed by replacing the integral inequality $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq 0$ with a stronger, but tractable, integral inequality over the space H in (7). This strategy is complementary to the approach followed in Section III, where we effectively replaced the space H with a tractable subspace S_N. In particular, we look for a lower bound $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq \mathcal{G}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$, where $\mathcal{G}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ is a functional whose nonnegativity over H can be enforced via a set of LMIs. Any $\boldsymbol{\gamma}$ such that $\mathcal{G}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq 0$ on H is then also feasible for (6), and the corresponding cost $\boldsymbol{c}^T \boldsymbol{\gamma}$ is an upper bound for the optimal value of (6).

A. Legendre Series Expansions

The key to constructing an inner approximation for the feasible set of problem (6) is to construct a functional $\mathcal{G}_{\boldsymbol{\gamma}} : H \to \mathbb{R}$, such that $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq \mathcal{G}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ for all $\boldsymbol{w} \in H$. To do this, we expand the components u and v of $\boldsymbol{w}$ (recall our simplifying restriction to the 2-D case) in terms of Legendre polynomials. That is, we write expansions such as

$$\partial^\alpha u = \sum_{n=0}^{\infty} \hat{u}_n^\alpha \mathcal{L}_n(x) \tag{14}$$

where $\mathcal{L}_n(x)$ is the Legendre polynomial of degree n and $\hat{u}_n^\alpha$ is the nth Legendre coefficient. Similar expressions can be written for v and its derivatives.

Legendre series expansions are useful because the Legendre polynomials are orthogonal on $[-1, 1]$, i.e., $\int_{-1}^{1} \mathcal{L}_m \, \mathcal{L}_n \, \mathrm{d}x = 0$ if $m \neq n$ [30]. This will enable us to enforce the nonnegativity of the functional $\mathcal{F}_{\boldsymbol{\gamma}}$ in (6) with a set of finite-dimensional, numerically tractable conditions. Note that although other polynomial basis functions, e.g., Chebyshev polynomials, may have more attractive numerical properties and may be more appropriate to implement the outer SDP relaxation of Theorem 1, they cannot be used here because they are only orthogonal with respect to a weighting function. A short introduction to Legendre polynomials, Legendre series and their properties is given in Appendix A; see [30]–[32] for a comprehensive treatment of the subject.

To avoid working with infinite series and to facilitate our analysis, we decompose (14) into a finite sum and a remainder function. More precisely, given an integer i we define the remainder function

$$U_i^\alpha(x) = \sum_{n=i+1}^{\infty} \hat{u}_n^\alpha \mathcal{L}_n(x). \tag{15}$$

Next, we choose an integer N, such that

$$N \geq d_F + k - 1 \tag{16}$$

where d_F is the degree of the polynomial matrix $\boldsymbol{F}$ defined in (5). For each $\alpha \in \{1, \ldots, k\}$, we decompose the Legendre expansion of $\partial^\alpha u$ as

$$\partial^\alpha u = \sum_{n=0}^{N+\alpha} \hat{u}_n^\alpha \, \mathcal{L}_n(x) + U_{N+\alpha}^\alpha(x). \tag{17}$$

For notational ease, we record the Legendre coefficients $\hat{u}_r^\alpha, \ldots, \hat{u}_s^\alpha$ for any two integers $0 \leq r \leq s$ in the vector

$$\hat{\boldsymbol{u}}_{[r,s]}^\alpha = [\, \hat{u}_r^\alpha, \ \ldots, \ \hat{u}_s^\alpha \,]^T \in \mathbb{R}^{s-r+1}. \tag{18}$$

For technical reasons that will be pointed out in Section IV-B, it will also be convenient to introduce an "extended" decomposition for the highest order derivative, $\partial^k u$. Specifically, let

$$M := N + 2k + d_F \tag{19}$$

and consider

$$\partial^k u = \sum_{n=0}^{M} \hat{u}_n^k \, \mathcal{L}_n(x) + U_M^k(x). \tag{20}$$

The following result, proven in Appendix B-B, relates the Legendre coefficients of $u, \partial u, \ldots, \partial^k u$.

Lemma 1: Let $u \in C^m([-1, 1])$ and its derivatives up to order $k \leq m - 1$ be expanded as in (17), and let M be as in (19). For any $\alpha \in \{1, \ldots, k\}$ and any two integers r, s with $0 \leq r \leq s \leq M + \alpha - k$, there exist matrices $\boldsymbol{B}_{[r,s]}^\alpha$ and $\boldsymbol{D}_{[r,s]}^\alpha$, such that

$$\hat{\boldsymbol{u}}_{[r,s]}^\alpha = \boldsymbol{B}_{[r,s]}^\alpha \, \mathcal{D}^{k-1} u|_{-1} + \boldsymbol{D}_{[r,s]}^\alpha \hat{\boldsymbol{u}}_{[0,M]}^k.$$

Furthermore, $\boldsymbol{B}_{[r,s]}^\alpha = \boldsymbol{0}$ if $r \geq k - \alpha$.

This lemma simply states that given the Legendre coefficients $\hat{u}_0^k, \ldots, \hat{u}_M^k$ of $\partial^k u$, the Legendre coefficients of all derivatives of order $\alpha < k$ can be computed uniquely if the boundary values $\mathcal{D}^{k-1} u|_{-1}$ are specified. These boundary values play the role of integration constants, and should be treated as variables until specific BCs are prescribed. Given an integer n, we therefore define the vector of variables

$$\check{\boldsymbol{u}}_n = \left[\left(\mathcal{D}^{k-1} u|_{-1} \right)^T, \ \hat{u}_0^k, \ \ldots, \ \hat{u}_n^k \right]^T \in \mathbb{R}^{k+n+1}. \tag{21}$$

The boundary values of u and its derivatives can also be represented in terms of our Legendre expansions. This is useful because the integral inequality in (6) is only required to hold for functions that satisfy prescribed BCs. The following result is proven in Appendix B-C.

Lemma 2: Let $u \in C^m([-1,1])$ and its derivatives up to order $k \leq m-1$ be expanded as in (17), and let $\mathcal{B}^{k-1}u \in \mathbb{R}^{2k}$ be defined according to (3). Moreover, let M be as in (19), and let $\check{\boldsymbol{u}}_M \in \mathbb{R}^{k+M+1}$ be defined according to (21). There exists a matrix $\boldsymbol{G}_M \in \mathbb{R}^{2k\times(k+M+1)}$, such that $\mathcal{B}^{k-1}u = \boldsymbol{G}_M\,\check{\boldsymbol{u}}_M$.

B. Legendre Expansion of $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$

Recalling the definition of $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$, we see from (6) that $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ is a sum of elementary terms of the form

$$\int_{-1}^{1} f\,\partial^\alpha u\,\partial^\beta v\,\mathrm{d}x \tag{22}$$

where $\alpha,\beta \in \{0,\ldots,k\}$. Here, $f = f(x;\boldsymbol{\gamma})$ denotes the appropriate entry of the integrand matrix $\boldsymbol{F}(x;\boldsymbol{\gamma})$ and, consequently, it is a polynomial of degree at most d_F whose coefficients are affine in $\boldsymbol{\gamma}$. We consider a term involving both components u and v of $\boldsymbol{w}$ for generality, but the following arguments also hold when $\partial^\alpha u\,\partial^\beta v$ is replaced with $\partial^\alpha u\,\partial^\beta u$ or $\partial^\alpha v\,\partial^\beta v$.

For each term of the form (22), we substitute $\partial^\alpha u$ and $\partial^\beta v$ with their decomposed Legendre expansions according to the following strategy.

1) If $\alpha \neq k$ or $\beta \neq k$, use (17).
2) If $\alpha = \beta = k$, use the "extended" decomposition (20).

The reasons for this choice will be explained in Remark 4, after Lemma 4. In either case, we can rewrite (22) as

$$\int_{-1}^{1} f\,\partial^\alpha u\,\partial^\beta v\,\mathrm{d}x = \mathcal{P}_{uv}^{\alpha\beta} + \mathcal{Q}_{uv}^{\alpha\beta} + \mathcal{R}_{uv}^{\alpha\beta} \tag{23}$$

where

$$\mathcal{P}_{uv}^{\alpha\beta} = \sum_{m=0}^{N_\alpha}\sum_{n=0}^{N_\beta} \hat{u}_m^\alpha\,\hat{v}_n^\beta \int_{-1}^{1} f\,\mathcal{L}_m\,\mathcal{L}_n\,\mathrm{d}x, \tag{24a}$$

$$\mathcal{Q}_{uv}^{\alpha\beta} = \sum_{n=0}^{N_\alpha} \hat{u}_n^\alpha \int_{-1}^{1} f\,\mathcal{L}_n\,V_{N_\beta}^\beta\,\mathrm{d}x + \sum_{n=0}^{N_\beta} \hat{v}_n^\beta \int_{-1}^{1} f\,\mathcal{L}_n\,U_{N_\alpha}^\alpha\,\mathrm{d}x, \tag{24b}$$

$$\mathcal{R}_{uv}^{\alpha\beta} = \int_{-1}^{1} f\,U_{N_\alpha}^\alpha\,V_{N_\beta}^\beta\,\mathrm{d}x. \tag{24c}$$

Here and in the following it should be understood that $N_\alpha = N+\alpha$ and $N_\beta = N+\beta$ if (17) is used to expand $\partial^\alpha u$ and $\partial^\beta v$, while $N_\alpha = N_\beta = M = N+2k+d_F$ if (20) is used.

The term $\mathcal{P}_{uv}^{\alpha\beta}$ is finite dimensional, and for any choice of $\alpha,\beta \in \{0,\ldots,k\}$ it can be rewritten as a symmetric quadratic form for the vectors $\hat{\boldsymbol{u}}_{[0,N_\alpha]}^\alpha$ and $\hat{\boldsymbol{v}}_{[0,N_\beta]}^\beta$. Recalling Lemma 1 and defining

$$\boldsymbol{\psi}_M := \begin{bmatrix} \check{\boldsymbol{u}}_M \\ \check{\boldsymbol{v}}_M \end{bmatrix} \in \mathbb{R}^{2(k+M+1)} \tag{25}$$

where $\check{\boldsymbol{u}}_M$ and $\check{\boldsymbol{v}}_M$ are as in (21), we arrive at the following result.

Lemma 3: Let $\mathcal{P}_{uv}^{\alpha\beta}$ be as in (24a) and $\boldsymbol{\psi}_M$ be defined according to (25). There exists a matrix $\boldsymbol{P}_{uv}^{\alpha\beta}(\boldsymbol{\gamma}) \in \mathbb{S}^{2(k+M+1)}$, whose entries are affine in $\boldsymbol{\gamma}$, such that

$$\mathcal{P}_{uv}^{\alpha\beta} = {\boldsymbol{\psi}_M}^T\,\boldsymbol{P}_{uv}^{\alpha\beta}(\boldsymbol{\gamma})\,\boldsymbol{\psi}_M.$$

The term $\mathcal{Q}_{uv}^{\alpha\beta}$ is less straightforward to handle, because it couples the first $N_\alpha+1$ and $N_\beta+1$ modes of $\partial^\alpha u$ and $\partial^\beta v$, respectively, to the remainder functions $V_{N_\beta}^\beta$ and $U_{N_\alpha}^\alpha$. We show in Appendix B-D that considering the extended decomposition (20) for the Legendre series of $\partial^k u$ and $\partial^k v$ enables us to write $\mathcal{Q}_{uv}^{\alpha\beta}$ as a finite-dimensional matrix quadratic form for the vector $\boldsymbol{\psi}_M$ if $\alpha \neq k$ or $\beta \neq k$. If $\alpha = \beta = k$, on the other hand, we cannot do the same unless f in (24b) is independent of x (in this case, the orthogonality of the Legendre polynomials and the remainder functions implies that $\mathcal{Q}_{uv}^{kk} = 0$). Instead, we estimate $\mathcal{Q}_{uv}^{kk}$ to decouple the remainder functions from the other terms.

To make these ideas more precise, let us introduce a family of "deflation" matrices $\boldsymbol{L}_n$, such that

$$\boldsymbol{L}_n\,\boldsymbol{\psi}_M = \begin{bmatrix} \hat{\boldsymbol{u}}_{[n,M]}^k \\ \hat{\boldsymbol{v}}_{[n,M]}^k \end{bmatrix}, \quad n \in \{0,\ldots,M\} \tag{26}$$

and $\boldsymbol{L}_n\,\boldsymbol{\psi}_M = \boldsymbol{0}$ if $n > M$. The existence of $\boldsymbol{L}_n$ follows from (18), (21), and (25). Moreover, given four integers $a \leq b$ and $c \leq d$, let $\boldsymbol{\Phi}_{[a,b]}^{[c,d]}$ be a $(b-a+1)\times(d-c+1)$ matrix whose ijth element is defined as

$$\left(\boldsymbol{\Phi}_{[a,b]}^{[c,d]}\right)_{ij} = \int_{-1}^{1} f\,\mathcal{L}_{m_i}\,\mathcal{L}_{n_j}\,\mathrm{d}x \tag{27}$$

where m_i and n_j are the ith and jth elements of the sequences $\{a,\ldots,b\}$ and $\{c,\ldots,d\}$. Note that, strictly speaking, $\boldsymbol{\Phi}_{[a,b]}^{[c,d]}$ depends on f, and its entries are affine on $\boldsymbol{\gamma}$. We do not indicate such dependences explicitly to avoid complicating our notation further. The following result is proven in Appendix B-D.

Lemma 4: Let $\mathcal{Q}_{uv}^{\alpha\beta}$ be as in (24b) and let d_F be the degree of $f(x;\boldsymbol{\gamma})$.

1) If $\alpha \neq k$ or $\beta \neq k$, there exists a matrix $\boldsymbol{Q}_{uv}^{\alpha\beta}(\boldsymbol{\gamma}) \in \mathbb{S}^{2(k+M+1)}$, whose entries are affine in $\boldsymbol{\gamma}$, such that
$$\mathcal{Q}_{uv}^{\alpha\beta} = {\boldsymbol{\psi}_M}^T\,\boldsymbol{Q}_{uv}^{\alpha\beta}(\boldsymbol{\gamma})\,\boldsymbol{\psi}_M.$$
2) If $\alpha = \beta = k$, let $\overline{M} := M+1-d_F$, define $\boldsymbol{\Delta} \in \mathbb{S}^{d_F}$ as
$$\boldsymbol{\Delta} := \mathrm{Diag}\left(\frac{2}{2(M+1)+1},\ \ldots,\ \frac{2}{2(M+d_F)+1}\right)$$
and define $\boldsymbol{Y}(\boldsymbol{\gamma}) \in \mathbb{R}^{2d_F\times 2d_F}$ as
$$\boldsymbol{Y}(\boldsymbol{\gamma}) := \frac{1}{2}\begin{bmatrix} \boldsymbol{0} & \boldsymbol{\Phi}_{[M+1-d_F,M]}^{[M+1,M+d_f]} \\ \boldsymbol{\Phi}_{[M+1,M+d_f]}^{[M+1-d_F,M]} & \boldsymbol{0} \end{bmatrix}.$$
Finally, let $\boldsymbol{Q}_{uv}^{kk} \in \mathbb{S}^{2d_F}$ and a diagonal matrix $\boldsymbol{\Sigma}_{uv}^{kk} \in \mathbb{S}^2$ satisfy the LMI
$$\boldsymbol{\Omega}(\boldsymbol{Q}_{uv}^{kk},\boldsymbol{\Sigma}_{uv}^{kk},\boldsymbol{\gamma}) := \begin{bmatrix} \boldsymbol{Q}_{uv}^{kk} & \boldsymbol{Y}(\boldsymbol{\gamma}) \\ \boldsymbol{Y}(\boldsymbol{\gamma})^T & \boldsymbol{\Sigma}_{uv}^{kk}\otimes\boldsymbol{\Delta} \end{bmatrix} \succeq 0 \tag{28}$$
where $\otimes$ is the usual Kronecker product. Then, $\mathcal{Q}_{uv}^{kk}$ can be bounded as in (29).

Remark 3: The LMI (28) was chosen, such that (29) shown at the bottom of this page (essentially its Schur complement condition) separates the contributions of $\boldsymbol{\psi}_M$, U_M^k, and V_M^k. As will be demonstrated in Section VI-C, inequality (29) is the main source of conservativeness. To make (29) as sharp as possible, we consider $\boldsymbol{Q}_{uv}^{kk}$ and $\boldsymbol{\Sigma}_{uv}^{kk}$ as auxiliary variables, to be determined subject to (28).

Remark 4: $\mathcal{Q}_{uv}^{\alpha\beta}$ can be represented exactly only if we consider all Legendre coefficients of $\partial^k u$, $\partial^k v$ up to order M explicitly: This is what motivates the use of the extended decomposition (20) for these functions. Moreover, note that instead of using the bound (29) we could write $\mathcal{Q}_{uv}^{kk}$ exactly in terms of $\boldsymbol{\psi}_{M+d_F}$, but this does not suit our aims because $\boldsymbol{\psi}_{M+d_F}$ is not decoupled from U_M^k, V_M^k (the Legendre coefficients $\hat{u}_{M+i}^k$, $1 \le i \le d_F$ appear in the definition of U_M^k).

Lemmas 3 and 4 show that $\mathcal{P}_{uv}^{\alpha\beta}$ and $\mathcal{Q}_{uv}^{\alpha\beta}$ can be expressed or bounded using $\boldsymbol{\psi}_M$, U_M^k, and V_M^k for any $\alpha, \beta \in \{0, \dots, k\}$. If $\alpha = \beta = k$, (24c) also depends on U_M^k and V_M^k. The following result, proven in Appendix B-E, shows that $\mathcal{R}_{uv}^{\alpha\beta}$ can be bounded using the same quantities when $\alpha \neq k$ or $\beta \neq k$.

Lemma 5: Suppose $\alpha \neq k$ or $\beta \neq k$, and let $\hat{\boldsymbol{f}}(\boldsymbol{\gamma}) = [\hat{f}_1(\boldsymbol{\gamma}), \dots, \hat{f}_{d_F}(\boldsymbol{\gamma})]^T$ be the vector of Legendre coefficients of the polynomial f. There exist a positive semidefinite matrix $\boldsymbol{R}_{uv}^{\alpha\beta} \in \mathbb{S}^{2(M+k+1)}$ with $\|\boldsymbol{R}_{uv}^{\alpha\beta}\|_F \sim N^{\alpha+\beta-2k-1}$ and a positive definite matrix $\boldsymbol{\Sigma}_{uv}^{\alpha\beta} \in \mathbb{S}^2$ with $\|\boldsymbol{\Sigma}_{uv}^{\alpha\beta}\|_F \sim N^{\alpha+\beta-2k}$, such that $\mathcal{R}_{uv}^{\alpha\beta}$ is bounded as in (30) shown at the bottom of this page.

Remark 5: The scaling of the Frobenius norms of $\boldsymbol{R}_{uv}^{\alpha\beta}$ and $\boldsymbol{\Sigma}_{uv}^{\alpha\beta}$ with N reflects the fact that the magnitude of $\mathcal{R}_{uv}^{\alpha\beta}$ diminishes to zero as N is raised. In contrast to Lemma 4, consequently, the conservativeness of the estimates in Lemma 5 can be reduced by simply increasing N.

C. Lower Bound for $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$

Let us now combine Lemmas 3–5 to find a lower bounding functional $\mathcal{G}_{\boldsymbol{\gamma}}$ for the integral functional $\mathcal{F}_{\boldsymbol{\gamma}}$ in (6). To account for the different cases in Lemma 4, we consider the contributions from terms with $\alpha = \beta = k$ first.

Let $\boldsymbol{S}(x;\boldsymbol{\gamma})$ be the symmetric matrix obtained from the rows and columns of the matrix $\boldsymbol{F}(x;\boldsymbol{\gamma})$ in (6) corresponding to the entries $\partial^k u$ and $\partial^k v$ of $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$. The contribution of the terms with $\alpha = \beta = k$ to $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ is

$$\int_{-1}^{1} \begin{bmatrix} \partial^k u \\ \partial^k v \end{bmatrix}^T \boldsymbol{S}(x;\boldsymbol{\gamma}) \begin{bmatrix} \partial^k u \\ \partial^k v \end{bmatrix} \mathrm{d}x.$$

It follows from Lemma 3 and part 2) of Lemma 4 that

$$\begin{aligned} \int_{-1}^{1} \begin{bmatrix} \partial^k u \\ \partial^k v \end{bmatrix}^T \boldsymbol{S} \begin{bmatrix} \partial^k u \\ \partial^k v \end{bmatrix} \mathrm{d}x \ge \boldsymbol{\psi}_M{}^T \Big[\boldsymbol{P}_{uu}^{kk} + 2\boldsymbol{P}_{uv}^{kk} + \boldsymbol{P}_{vv}^{kk} \\ - \boldsymbol{L}_{\overline{M}}{}^T \left(\boldsymbol{Q}_{uu}^{kk} + 2\boldsymbol{Q}_{uv}^{kk} + \boldsymbol{Q}_{vv}^{kk} \right) \boldsymbol{L}_{\overline{M}} \Big] \boldsymbol{\psi}_M \\ + \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \left[\boldsymbol{S} - \boldsymbol{\Sigma}_{uu}^{kk} - 2\boldsymbol{\Sigma}_{uv}^{kk} - \boldsymbol{\Sigma}_{vv}^{kk} \right] \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} \mathrm{d}x \end{aligned} \tag{31}$$

where the auxiliary variables $\boldsymbol{Q}_{uu}^{kk}$, $\boldsymbol{\Sigma}_{uu}^{kk}$, $\boldsymbol{Q}_{uv}^{kk}$, $\boldsymbol{\Sigma}_{uv}^{kk}$, $\boldsymbol{Q}_{vv}^{kk}$, and $\boldsymbol{\Sigma}_{vv}^{kk}$ must satisfy three LMIs defined as in (28). For notational convenience, we let

$$\mathcal{Y} = \left\{ \boldsymbol{Q}_{uu}^{kk}, \boldsymbol{\Sigma}_{uu}^{kk}, \boldsymbol{Q}_{uv}^{kk}, \boldsymbol{\Sigma}_{uv}^{kk}, \boldsymbol{Q}_{vv}^{kk}, \boldsymbol{\Sigma}_{vv}^{kk} \right\} \tag{32}$$

be the list of all auxiliary variables, and we combine the three LMIs they must satisfy into the equivalent block-diagonal LMI (33) shown at the bottom of this page.

All terms contributing to $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ with $\alpha \neq k$ or $\beta \neq k$ can instead be lower bounded using Lemmas 3–5 to obtain expressions such as (34) shown at the bottom of this page.

From (31) and (34) we conclude that it is possible to construct a matrix $\boldsymbol{Q}_M = \boldsymbol{Q}_M(\boldsymbol{\gamma}, \mathcal{Y}) \in \mathbb{S}^{2(k+M+1)}$ and a positive definite matrix $\boldsymbol{\Sigma}_M = \boldsymbol{\Sigma}_M(\boldsymbol{\gamma}, \mathcal{Y}) \in \mathbb{S}^2$, such that for all $\boldsymbol{w}$

$$\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \ge \boldsymbol{\psi}_M{}^T \boldsymbol{Q}_M \boldsymbol{\psi}_M + \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T (\boldsymbol{S} - \boldsymbol{\Sigma}_M) \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} \mathrm{d}x. \tag{35}$$

Note that $\boldsymbol{Q}_M$ and $\boldsymbol{\Sigma}_M$ are affine with respect to the variables listed in $\mathcal{Y}$ but *not* in $\boldsymbol{\gamma}$, because Lemma 5 introduces absolute values of linear functions of $\boldsymbol{\gamma}$.

D. Projection Onto the BCs

The lower bound (35) holds for any continuously differentiable function $\boldsymbol{w}$, irrespective of whether it satisfies the BCs prescribed on H. Recalling (7), these are given by the set of p

$$\mathcal{Q}_{uv}^{kk} \ge -\boldsymbol{\psi}_M{}^T \left(\boldsymbol{L}_{\overline{M}}{}^T \boldsymbol{Q}_{uv}^{kk} \boldsymbol{L}_{\overline{M}} \right) \boldsymbol{\psi}_M - \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \boldsymbol{\Sigma}_{uv}^{kk} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} \mathrm{d}x. \tag{29}$$

$$\left| \mathcal{R}_{uv}^{\alpha\beta} \right| \le \|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 \, \boldsymbol{\psi}_M{}^T \boldsymbol{R}_{uv}^{\alpha\beta} \boldsymbol{\psi}_M + \|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \boldsymbol{\Sigma}_{uv}^{\alpha\beta} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} \mathrm{d}x. \tag{30}$$

$$\overline{\boldsymbol{\Omega}}(\boldsymbol{\gamma}, \mathcal{Y}) := \begin{bmatrix} \boldsymbol{\Omega}(\boldsymbol{Q}_{uu}^{kk}, \boldsymbol{\Sigma}_{uu}^{kk}, \boldsymbol{\gamma}) & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega}(\boldsymbol{Q}_{uv}^{kk}, \boldsymbol{\Sigma}_{uv}^{kk}, \boldsymbol{\gamma}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\Omega}(\boldsymbol{Q}_{vv}^{kk}, \boldsymbol{\Sigma}_{vv}^{kk}, \boldsymbol{\gamma}) \end{bmatrix} \succeq 0. \tag{33}$$

$$\int_{-1}^{1} f \, \partial^\alpha u \, \partial^\beta v \, \mathrm{d}x \ge \boldsymbol{\psi}_M{}^T \left(\boldsymbol{P}_{uv}^{\alpha\beta} + \boldsymbol{Q}_{uv}^{\alpha\beta} - \|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 \boldsymbol{R}_{uv}^{\alpha\beta} \right) \boldsymbol{\psi}_M - \|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \boldsymbol{\Sigma}_{uv}^{\alpha\beta} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} \mathrm{d}x. \tag{34}$$

homogeneous equations

$$\boldsymbol{A}\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} = \boldsymbol{0}. \tag{36}$$

To enforce as many BCs as possible in (35) and to sharpen the lower bound over the space H, we need to rewrite (36) in terms of our Legendre expansions. We begin by introducing a permutation matrix $\boldsymbol{P}$, such that

$$\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} = \boldsymbol{P}\begin{bmatrix}\mathcal{B}^{\boldsymbol{k-1}}\boldsymbol{w}\\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}\end{bmatrix} \tag{37}$$

so (36) becomes

$$\boldsymbol{A}\boldsymbol{P}\begin{bmatrix}\mathcal{B}^{\boldsymbol{k-1}}\boldsymbol{w}\\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}\end{bmatrix} = \boldsymbol{0}. \tag{38}$$

A straightforward corollary of Lemma 2 and (25) is that there exists a matrix $\boldsymbol{J}$, such that $\mathcal{B}^{\boldsymbol{k-1}}\boldsymbol{w} = \boldsymbol{J}\boldsymbol{\psi}_M$. Then, (38) can be rewritten as

$$\boldsymbol{K}\begin{bmatrix}\boldsymbol{\psi}_M\\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}\end{bmatrix} = \boldsymbol{0}, \quad \boldsymbol{K} := \boldsymbol{A}\boldsymbol{P}\begin{bmatrix}\boldsymbol{J} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}. \tag{39}$$

From (39) we see that any admissible vector $\boldsymbol{\psi}_M$ can be written in the form

$$\boldsymbol{\psi}_M = \boldsymbol{\Pi}_M\boldsymbol{\zeta} \tag{40}$$

for some $\boldsymbol{\zeta} \in \mathbb{R}^{\dim[\mathcal{N}(\boldsymbol{K})]}$, where $\boldsymbol{\Pi}_M$ is a computable projection matrix. Note that in general $\boldsymbol{\Pi}_M$ may have linearly dependent columns and so it may be further simplified. This makes no difference to the following discussion, and we omit the details to streamline the presentation. Substituting (40) into (35), we conclude that when (33) holds, $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ is lower bounded over the space H in (7) as

$$\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq \boldsymbol{\zeta}^T\,\boldsymbol{\Pi}_M^T\boldsymbol{Q}_M\boldsymbol{\Pi}_M\,\boldsymbol{\zeta} + \int_{-1}^{1}\begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix}^T[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M]\begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix}\mathrm{d}x. \tag{41}$$

From (39) and (40) it is also possible to formulate a set of BCs that further restrict the choice for U_M^k and V_M^k. Moreover, recall that the remainder functions U_M^k and V_M^k should be orthogonal to all Legendre polynomials of degree less than or equal to M. However, it is not currently clear to the authors how these two constraints can be enforced explicitly in (41) to obtain a stronger, but still useful, lower bound on $\mathcal{F}_{\boldsymbol{\gamma}}$. Consequently, we choose to simply drop them and let U_M^k and V_M^k be arbitrary functions.

E. Formulating an Inner SDP Relaxation

The integral inequality in (6) is satisfied if the right-hand side of (41) is nonnegative for all $\boldsymbol{\zeta}$ and all functions U_M^k and V_M^k. Recalling that (41) is valid only if (33) holds, we have the following result.

Proposition 1: Let $M = M(N)$ be as in (19) for any integer N, and let $\mathcal{Y}$ be as in (32). The set $T_N^{\text{in}} \subset \mathbb{R}^s$ of values $\boldsymbol{\gamma} \in \mathbb{R}^s$ for which there exist $\mathcal{Y}$, such that

$$\overline{\boldsymbol{\Omega}}(\boldsymbol{\gamma};\mathcal{Y}) \succeq 0, \tag{42a}$$

$$\boldsymbol{\Pi}_M^T\,\boldsymbol{Q}_M(\boldsymbol{\gamma},\mathcal{Y})\,\boldsymbol{\Pi}_M \succeq 0, \tag{42b}$$

$$\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y}) \succeq 0, \quad \forall x \in [-1,1] \tag{42c}$$

is an inner approximation of the feasible set T of (6), i.e., $T_N^{\text{in}} \subset T$.

Conditions (42b) and (42c) are only sufficient, not necessary, to make the right-hand side of (41) nonnegative: They do not take into account the boundary and orthogonality conditions on the remainder functions mentioned at the end of Section IV-D. However, they are useful because they can be turned into tractable constraints. For example, (42b) is not an LMI because $\boldsymbol{Q}_M(\boldsymbol{\gamma},\mathcal{Y})$ depends on absolute values of the Legendre coefficients of the entries of the matrix $\boldsymbol{F}(x;\boldsymbol{\gamma})$ in (5) as a consequence of Lemma 5. However, it can readily be recast as one by replacing each of these absolute values, say $|\hat{f}_n(\boldsymbol{\gamma})|$, with a slack variable t subject to the additional linear constraints $-t \leq \hat{f}_n(\boldsymbol{\gamma}) \leq t$ [33]. Moreover, (42c) is an LMI if the matrix $\boldsymbol{S}(x;\boldsymbol{\gamma})$ is independent of x, which is true in many interesting and nontrivial cases, such as our motivating example in Section II. Otherwise, (42c) is equivalent to the polynomial inequality

$$\boldsymbol{z}^T[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y})]\boldsymbol{z} \geq 0,\ \forall(x,\boldsymbol{z}) \in [-1,1]\times\mathbb{R}^2. \tag{43}$$

Although checking a polynomial inequality is generally NP-hard (see [34, Section 2.1] and references therein), we can turn (43) into an LMI plus linear equality constraints by an SOS relaxation [34]. Using the so-called $\mathcal{S}$-procedure [35], we introduce a tunable symmetric polynomial matrix $\boldsymbol{T}(x) \in \mathbb{S}^2$ and require that the multivariate polynomials

$$p_1 := \boldsymbol{z}^T[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y}) - (1-x^2)\boldsymbol{T}(x)]\boldsymbol{z},$$
$$p_2 := \boldsymbol{z}^T\boldsymbol{T}(x)\boldsymbol{z}$$

are SOS; it is not difficult to see that this implies (43).

An upper bound for the optimal value of (6), as well as a feasible point that achieves it, can therefore be found by solving an SDP.

Theorem 2: Let $M = M(N)$ be defined as in (19) for any integer N, let $\mathcal{Y}$ be as in (32), and let $\boldsymbol{T}(x) \in \mathbb{S}^2$ be a tunable polynomial matrix. The optimal value of the SDP

$$\begin{aligned}\min_{\boldsymbol{\gamma},\mathcal{Y},\boldsymbol{T}(x)}\quad & \boldsymbol{c}^T\boldsymbol{\gamma},\\ \text{s.t.}\quad & \overline{\boldsymbol{\Omega}}(\boldsymbol{\gamma};\mathcal{Y}) \succeq 0,\\ & \boldsymbol{\Pi}_M^T\,\boldsymbol{Q}_M(\boldsymbol{\gamma},\mathcal{Y})\,\boldsymbol{\Pi}_M \succeq 0,\\ & \boldsymbol{z}^T\left[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y}) - (1-x^2)\boldsymbol{T}(x)\right]\boldsymbol{z} \quad \text{is SOS},\\ & \boldsymbol{z}^T\boldsymbol{T}(x)\boldsymbol{z} \quad \text{is SOS}\end{aligned} \tag{44}$$

is an upper bound for the optimal value of (6). Moreover, if a minimizer $\boldsymbol{\gamma}_N^\star$ exists in (44), it is a feasible point for (6).

Remark 6: In contrast to our results for the outer SDP relaxations of Section III, we cannot prove that the optimal value of (44) converges to that of the original problem as N is increased, nor that it is nonincreasing. In fact, without further assumptions on the functional $\mathcal{F}_{\boldsymbol{\gamma}}$ in (6), it is possible that (44) is always infeasible even if (6) is feasible. To see this, recall that the matrix $\boldsymbol{\Sigma}_M$ is positive definite, so (42c) and its corresponding SOS relaxation are feasible only if $\boldsymbol{S}(x;\boldsymbol{\gamma})$ can be made sufficiently positive definite for all $x \in [-1,1]$. An example for which this does not happen is the integral inequality

$$\int_{-1}^{1}\left[x^2(\partial u)^2 + (\partial v)^2 - \gamma uv\right]\mathrm{d}x \geq 0 \tag{45}$$

where u and v are subject to the Dirichlet BCs $u|_{-1} = u|_1 = v|_{-1} = v|_1 = 0$. This inequality is clearly feasible for $\gamma = 0$. Yet, (44) is infeasible for any N because $\boldsymbol{S}(x;\boldsymbol{\gamma}) = \left[\begin{smallmatrix} x^2 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ is not positive definite at $x = 0$. In fact, for this particular example, *any* approach requiring estimates of tail terms of series expansions will necessarily be ineffective. In contrast, with the SOS method of [3] we could establish that (45) is feasible for $|\gamma| \leq 2.2$ at least. With the exception of such pathological cases, however, our inner SDP relaxations are observed to work well in practice, and we demonstrate this in Section VI. This suggests that it may be possible to formulate precise conditions under which our inner SDP relaxations are feasible, and even converge to the original optimization problem. We leave this task to future research.

V. Extensions

A. Inequalities With Explicit Dependence on Boundary Values

In the applications we have in mind, the integral inequality constraint in (6) is derived from a weak formulation of a PDE, after integrating some terms by parts. Occasionally, the BCs are such that the boundary terms from such integrations by parts do not vanish (we will give an example in Section VI-B).

This motivates us to extend our results to quadratic homogeneous functionals that depend explicitly on the boundary values $\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}$, such as (46) shown at the bottom of this page, where $\boldsymbol{F}_{\text{int}}$, $\boldsymbol{F}_{\text{mix}}$, and $\boldsymbol{F}_{\text{bnd}}$ are matrices of polynomials of degree at most d_F of the form (5). Note that $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ in (46) reduces to the functional in (6) if $\boldsymbol{F}_{\text{mix}} = \boldsymbol{0}$, $\boldsymbol{F}_{\text{bnd}} = \boldsymbol{0}$, and $\boldsymbol{F}_{\text{int}} = \boldsymbol{F}$.

The extension of Theorem 1 is obvious, because the boundary values of polynomial functions are easily given in terms of the polynomial coefficients.

To extend Proposition 1 and Theorem 2, we recall the definition of the permutation matrix $\boldsymbol{P}$ in (37). Upon integrating the known matrix $\boldsymbol{P}^T\boldsymbol{F}_{\text{bnd}}(x;\boldsymbol{\gamma})\boldsymbol{P}$, it follows from (25) and Lemma 2 that there exists a symmetric matrix $\boldsymbol{Q}_M^{\text{bnd}}(\boldsymbol{\gamma})$, such that (47) holds.

Moreover, let $\boldsymbol{g}(x;\boldsymbol{\gamma})$ be the column of the matrix $\boldsymbol{P}^T\boldsymbol{F}_{\text{mix}}(x;\boldsymbol{\gamma})$ corresponding to the entry $\partial^\alpha u$ of $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$. Each element $g_i(x;\boldsymbol{\gamma})$ is a polynomial of degree at most d_F, written in the Legendre basis with coefficients $\hat{g}_{i,0}(\boldsymbol{\gamma}), \ldots, \hat{g}_{i,d_F}(\boldsymbol{\gamma})$. Recalling from (16) that we have decomposed the Legendre expansion of $\partial^\alpha u$ with the truncation parameter $N \geq d_F + k - 1$, we can write (48) shown at the bottom of this page. With the help of Lemma 1, (25), and Lemma 2 it is then possible to find a matrix $\boldsymbol{Q}_M^{\text{mix}}(\boldsymbol{\gamma})$ that satisfies (49) shown at the bottom of the this page.

Note that (47) shown at the bottom of the this page, and (49) are exact formulae, and no approximation is made. Combining these results with (35) we conclude that there is a symmetric matrix $\boldsymbol{Q}_M^{\text{tot}} = \boldsymbol{Q}_M^{\text{tot}}(\boldsymbol{\gamma}, \mathcal{Y})$ such that (50) shown at the bottom of this page holds. Finally, (39) implies that we can write

$$\begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix} = \boldsymbol{\Lambda}\boldsymbol{\zeta}$$

for some $\boldsymbol{\zeta} \in \mathbb{R}^{\dim[\mathcal{N}(\boldsymbol{K})]}$, where the projection matrix $\boldsymbol{\Lambda}$ satisfies $\mathcal{R}(\boldsymbol{\Lambda}) = \mathcal{N}(\boldsymbol{K})$, and we conclude that Proposition 1 and Theorem 2 are true when we replace (42b) and the corresponding constraint in (44) with $\boldsymbol{\Lambda}^T\boldsymbol{Q}_M^{\text{tot}}(\boldsymbol{\gamma})\boldsymbol{\Lambda} \succeq 0$.

B. Higher Dimensional Function Spaces and Generic Multi-Index Derivatives

Theorems 1 and 2 were derived with the assumption that $\boldsymbol{w} \in C^m([-1,1],\mathbb{R}^2)$ and for the particular multi-indices $\boldsymbol{k} = [k,k]$, $\boldsymbol{l} = [l,l]$. All our statements, including the extensions discussed in Section V-A, hold also when we let $\boldsymbol{w} \in C^m([-1,1],\mathbb{R}^q)$ with $q \geq 1$ and when $\boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}^q$ are generic multi-indices, as long as they satisfy (4a) and (4b).

In particular, all our proofs extend verbatim by simply identifying the functions u, v used throughout Sections III and IV with any two components w_i, w_j of $\boldsymbol{w}$ if the q-dimensional multi-indices $\boldsymbol{k}$ and $\boldsymbol{l}$ are uniform, i.e., $\boldsymbol{k} = [k, \ldots, k]$ and $\boldsymbol{l} = [l, \ldots, l]$. The extension to nonuniform multi-indices $\boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}^q$ requires only minor modifications, the details are left to the interested reader.

VI. Computational Experiments With QUINOPT

In this section, we apply our techniques to solve some problems arising from the analysis of PDEs. To aid the formulation of our SDP relaxations, we have developed QUINOPT, an open-source add-on for the MATLAB optimization toolbox YALMIP [27], [28]. QUINOPT uses the Legendre polynomial basis for the outer SDP relaxations of Section III, because the

$$\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \int_{-1}^{1} \left[\left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{bnd}}(x;\boldsymbol{\gamma})\, \mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} + \left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{mix}}(x;\boldsymbol{\gamma})\, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w} + \left(\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{int}}(x;\boldsymbol{\gamma})\, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w} \right] \mathrm{d}x, \tag{46}$$

$$\int_{-1}^{1} \left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{bnd}}(x;\boldsymbol{\gamma})\, \mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\, \mathrm{d}x = \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \boldsymbol{Q}_M^{\text{bnd}}(\boldsymbol{\gamma}) \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}. \tag{47}$$

$$\int_{-1}^{1} g_i(x;\boldsymbol{\gamma})\partial^\alpha u\, \mathrm{d}x = \sum_{m=0}^{d_F}\sum_{n=0}^{\infty} \hat{g}_{i,m}(\boldsymbol{\gamma})\hat{u}_n^\alpha \int_{-1}^{1} \mathcal{L}_m\, \mathcal{L}_n\, \mathrm{d}x = \left[2\hat{g}_{i,0}(\boldsymbol{\gamma}), \quad \frac{2\hat{g}_{i,1}(\boldsymbol{\gamma})}{3}, \quad \ldots, \quad \frac{2\hat{g}_{i,d_F}(\boldsymbol{\gamma})}{2d_F+1} \right] \hat{\boldsymbol{u}}^\alpha_{[0,d_F]}. \tag{48}$$

$$\int_{-1}^{1} \left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{mix}}(x;\boldsymbol{\gamma})\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\, \mathrm{d}x = \begin{bmatrix} \mathcal{B}^{\boldsymbol{k}-\boldsymbol{1}}\boldsymbol{w} \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \int_{-1}^{1} \boldsymbol{P}^T \boldsymbol{F}_{\text{mix}}(x;\boldsymbol{\gamma})\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\, \mathrm{d}x = \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \boldsymbol{Q}_M^{\text{mix}}(\boldsymbol{\gamma})\, \boldsymbol{\psi}_M. \tag{49}$$

$$\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \boldsymbol{Q}_M^{\text{tot}} \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix} + \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T [\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M] \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} \mathrm{d}x. \tag{50}$$

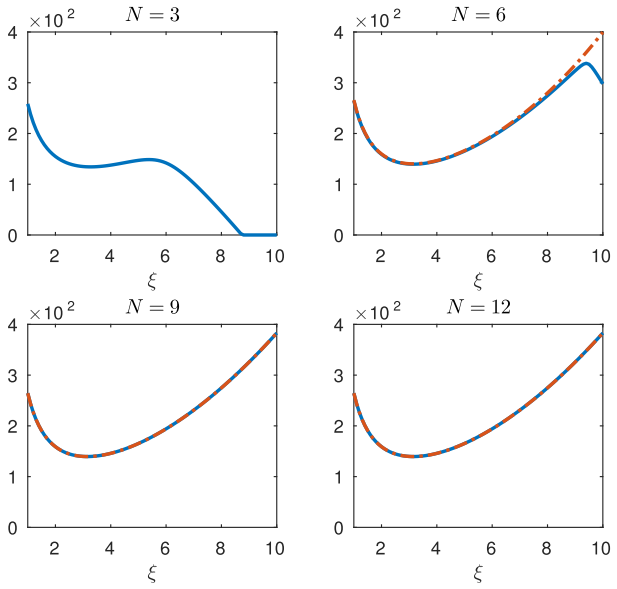


Fig. 2. (Color online) Upper (dotted–dashed line) and lower (solid line) bounds on the optimal value of (8) as a function of ξ for different values of N. The upper bound for $N = 3$ is infinite and so it is not plotted. The bounds are indistinguishable for $N = 9$ and $N = 12$.

orthogonality of the Legendre polynomials promotes sparsity of the SDP data. QUINOPT and the scripts used to produce the results in the following sections can be downloaded from https://github.com/aeroimperial-optimization/QUINOPT.

Our experiments were run on a PC with a 3.40 GHz Intel Core i7-4770 CPU and 16 Gb of RAM, using MOSEK [36] to solve our SDP relaxations.

A. Motivating Example: Stability of a Stress-Driven Shear Flow

Consider our motivating example of Section II-A, regarding the stability of a flow driven by a shear stress of magnitude 0.5γ. An ad hoc inner SDP relaxation was proposed and solved in [26]; here, we replicate those results using our general-purpose toolbox QUINOPT. Since we minimize the *negative* of γ in (8), the inner and outer SDPs (44) and (13) give, respectively, lower and upper bounds for the stress $\gamma_{\rm cr}$ at which the flow is no longer provably stable.

Fig. 2 shows the upper and lower bounds for $\gamma_{\rm cr}$ as a function of the wave number ξ, a parameter in (8), computed for four different values of the Legendre series truncation parameter N. No upper bound curve is plotted for $N = 3$ because in this case only the zero polynomial satisfies the BCs in (9), and (13) reduces to an unconstrained minimization problem yielding an infinite upper bound. More detailed numerical results, CPU times, and the number of primal and dual variables in the SDP relaxations (denoted n and m, respectively) are reported in Table I for wave number parameters $\xi = 3$ and $\xi = 9$. The table also gives results obtained when the SOS relaxations of [3] are formulated using polynomials of degree d.

Our results show that within the tested range of ξ, the upper and lower bounds converge to each other at relatively small values of N (three decimal places for $N = 12$ for both cases reported in Table I). This means that we can bound $\gamma_{\rm cr}$ accurately and extremely efficiently using our techniques, and that the inner SDP relaxations converge to the full problem (8) despite our inability to provide a proof of this fact in general (cf. Remark 6). Finally, note that our techniques significantly outperform the SOS method of [3] in terms of computational cost and quality of the lower bound.

B. Stability of a System of Coupled PDEs

Let $\boldsymbol{w} = [u(t,x),\, v(t,x)]^T$ and consider the system of PDEs

$$\partial_t \boldsymbol{w} = \gamma \partial_x^2 \boldsymbol{w} + \boldsymbol{A}\boldsymbol{w}, \qquad \boldsymbol{A} = \begin{bmatrix} 1 & 1.5 \\ 5 & 0.2 \end{bmatrix} \tag{51}$$

over the domain $[0,1]$, subject to the BCs $u|_0 = u|_1 = v|_0 = v|_1 = 0$. This system was studied in [37, Section V-D] with the equivalent parametrization $\gamma = R^{-1}$. The stabilizing effect of the diffusive term $\gamma\partial_x^2\boldsymbol{w}$ decreases with γ, until the equilibrium solution $[u,v]^T = [0,0]^T$ becomes unstable. It can be shown that the amplitude of infinitesimal sinusoidal perturbations to the zero solution grows exponentially in time if $\gamma < \gamma_{\rm cr} = 0.3412$. Since the system is linear, it is stable with respect to finite-amplitude perturbations for all $\gamma \geq \gamma_{\rm cr}$.

Following [37], we try to establish the stability of the system with respect to arbitrary perturbations by considering Lyapunov functionals of the form

$$\mathcal{V}(t) = \frac{1}{2}\int_0^1 \boldsymbol{w}^T \boldsymbol{P}(x)\boldsymbol{w}\,\mathrm{d}x \tag{52}$$

where $\boldsymbol{P}(x)$ is a tunable polynomial matrix of given degree d_P, such that $\mathcal{V}(t) \geq c\,\|\boldsymbol{w}\|_2^2$ for some $c > 0$ and $-\frac{\mathrm{d}\mathcal{V}}{\mathrm{d}t} \geq 0$. Note that since $\boldsymbol{P}(x)$ can always be rescaled by c without changing the sign of the inequalities, we may fix $c = 1$.

Using (51) to compute $\frac{\mathrm{d}\mathcal{V}}{\mathrm{d}t}$, we find that the critical value of γ at which (52) stops being a valid Lyapunov function for a given degree d_P is given by

$$\begin{aligned} \min_{\gamma,\boldsymbol{P}(x)} \quad & \gamma \\ \text{s.t.} \quad & \int_0^1 \boldsymbol{w}^T\left[\boldsymbol{P}(x) - \boldsymbol{I}\right]\boldsymbol{w}\,\mathrm{d}x \geq 0, \\ & \int_0^1 \boldsymbol{w}^T\boldsymbol{P}(x)(-\gamma\partial_x^2\boldsymbol{w} - \boldsymbol{A}\boldsymbol{w})\,\mathrm{d}x \geq 0. \end{aligned} \tag{53}$$

Note that although the system state $\boldsymbol{w}$ is a function of time, the integral inequalities above are imposed *pointwise in time*. Therefore, the time dependence can be formally dropped, and (53) is in the form (6) with two integral inequalities.

Since the optimization variables are γ and the coefficients of the entries of $\boldsymbol{P}(x)$, the problem is not jointly convex in γ and $\boldsymbol{P}$, and we cannot minimize γ directly. Instead, we fix a trial value for γ and check whether a feasible $\boldsymbol{P}(x)$ of degree d_P exists. The optimal γ for (53), which must be finite because the system is linearly unstable when γ is sufficiently small, is then given by the value at which a feasible $\boldsymbol{P}(x)$ ceases to exist, and it can be determined with a simple bisection procedure.

Before deriving our SDP relaxations, we need to rescale the domain of integration for the constraints in (53) to $[-1,1]$. Moreover, in light of Remark 6, the second integral inequality should be integrated by parts to prevent the inner SDP relaxation from being infeasible. Both tasks (rescaling and integration by parts) are performed automatically by QUINOPT. We also note that

TABLE I
UPPER AND LOWER BOUNDS ON $\gamma_{\rm cr}$ (DENOTED LB AND UB), CPU TIMES (IN SECONDS) AND PRIMAL–DUAL PROBLEM DIMENSIONS (n AND m) FOR THE OUTER AND INNER RELAXATIONS OF PROBLEM (8) FROM QUINOPT, AS A FUNCTION OF THE LEGENDRE TRUNCATION PARAMETER N

	QUINOPT, outer					QUINOPT, inner					SOS [3]				
	N	n	m	UB	t	N	n	m	LB	t	d	n	m	LB	t
$\xi = 3$	3	0	1	+INF	0.03	3	202	2	134.8594	0.09	4	805	230	79.4435	0.19
	6	36	1	140.4087	0.04	6	406	2	139.7656	0.10	8	2349	454	119.8619	0.28
	9	144	1	139.7701	0.06	9	683	2	139.7700	0.08	16	7789	902	130.5796	0.68
	12	324	1	139.7700	0.05	12	1030	2	139.7700	0.10	32	28077	1798	134.4737	3.16
$\xi = 9$	3	0	1	+INF	0.03	3	202	2	0.0000	0.08	4	805	230	285.9021	0.18
	6	36	1	335.1022	0.04	6	406	2	323.5764	0.08	8	2349	454	314.1146	0.27
	9	144	1	325.6764	0.05	9	683	2	325.6449	0.09	16	7789	902	321.2403	0.65
	12	324	1	325.6455	0.05	12	1030	2	325.6453	0.10	32	28077	1798	323.1421	2.98

Results for the inner SOS relaxation of [3] using polynomials of degree d are shown for comparison.

TABLE II
UPPER AND LOWER BOUNDS (UB AND LB) FOR THE OPTIMAL SOLUTION OF (53) FOR DIFFERENT CHOICES OF d_P, AND FOR THE CASE $\boldsymbol{P}(x) = \boldsymbol{I}$

d_P	UB from [2]	UB	$t_{\rm UB}$	LB	$t_{\rm LB}$
$\boldsymbol{P}(x) = \boldsymbol{I}$	5	0.3925	0.26	0.3925	0.09
0	3.3333	0.3412	0.14	0.3412	0.12
2	0.5882	0.3412	1.32	0.3412	0.99
4	0.4347	0.3412	1.57	0.3412	1.07
6	0.4166	0.3412	1.82	0.3412	1.18

Also reported are the average CPU times $t_{\rm UB}$ and $t_{\rm LB}$ to solve each feasibility problem in the bisection procedure to compute the upper/lower bounds UB and LB (to minimize γ in the case $\boldsymbol{P}(x) = \boldsymbol{I}$), and upper bounds from [37].

TABLE III
CPU TIME (IN SECONDS) FOR THE COMPUTATION OF THE SETS $T_N^{\rm out}$, $T_N^{\rm in}$, AND $T_N^{\rm sos}$ AS A FUNCTION OF N USING MOSEK AND SDPT3

	MOSEK			SDPT3		
N	$T_N^{\rm out}$	$T_N^{\rm in}$	$T_N^{\rm sos}$	$T_N^{\rm out}$	$T_N^{\rm in}$	$T_N^{\rm sos}$
2	0.81	1.84	1.58	10.5	25.4	21.5
4	0.95	2.36	4.10	12.6	30.9	45.1
8	1.66	4.99	19.2	14.5	46.4	176
16	4.82	6.80	346	21.1	58.2	600
24	9.36	9.11	2100	25.5	59.0	10500
32	17.7	14.6	6220	35.5	71.5	117000

after rescaling and integration by parts, the second integral inequality in (53) depends explicitly on the unspecified boundary values $\partial_x u|_{\pm 1}$ and $\partial_x v|_{\pm 1}$, making the extensions discussed in Section V-A necessary.

Table II shows upper and lower bounds for the optimal solution of (53) as a function of the degree d_P of $\boldsymbol{P}(x)$, obtained by applying the bisection procedure described above to the SDPs (44) and (13), respectively. We also show results for the particular choice $\boldsymbol{P}(x) = \boldsymbol{I}$, corresponding to the classical approach of taking the energy of the system as the candidate Lyapunov function; in this case, a direct minimization over γ could be performed. In all computations we fixed the Legendre series truncation parameter to $N = 10$ and the degree of the matrix $\boldsymbol{T}(x)$ in (44) to 6, which gives well-converged results. Table II also reports the average CPU time taken by QUINOPT to set up and solve each feasibility problem in our bisection procedure (to minimize γ when we fixed $\boldsymbol{P}(x) = \boldsymbol{I}$).

Our results show that stability can be established up to the known critical value $\gamma_{\rm cr} = 0.3412$ for all choices of d_P, with the exception of the classical energy Lyapunov function. This drastically improves the conservative results obtained with the SOS method in [37] for the same problem, also reported in Table II (the original results are for a parameter $R = \gamma^{-1}$ and have been adapted). Our results demonstrate that our SDP relaxations accurately approximate (53). This is particularly significant for the inner SDPs, which rely on typically conservative estimates and for which we cannot prove convergence.

C. Feasible Set Approximation

In this final example, we consider the problem of computing the entire feasible set of the integral inequality

$$\int_{-1}^{1} \left[(\partial u)^2 + (\partial v)^2 + \gamma_1 x^2 \partial u\, \partial v + 2\gamma_2 uv \right] {\rm d}x \geq 0 \quad (54)$$

where u and v are subject to the Dirichlet BCs $u|_{-1} = 0$, $u|_1 = 0$, $v|_{-1} = 0$, $v|_1 = 0$. This inequality does not arise from a particular PDE, but has been constructed *ad hoc* to illustrate some subtle properties of our SDP relaxations and highlight the main sources of conservativeness.

Outer and inner approximation sets $T_N^{\rm out}$ and $T_N^{\rm in}$ can be found using (13) and (44), respectively. In particular, we compute the boundaries of $T_N^{\rm out}$ and $T_N^{\rm in}$ by optimizing the objective function $\gamma_1 \sin\theta + \gamma_2 \cos\theta$ for 300 equispaced values of $\theta \in [0, 2\pi]$. When solving (44), we fix the degree of the tunable polynomial matrix $\boldsymbol{T}(x)$ to the smallest of $N - 2$ and 6; our results do not improve when this is increased. Inner approximation sets $T_N^{\rm sos}$ can be computed in a similar way using the SOS method of [3] with polynomials of degree N.

The CPU time required to compute $T_N^{\rm out}$, $T_N^{\rm in}$, and $T_N^{\rm sos}$ is shown in Table III for six values of N and two SDP solvers, MOSEK [36] and SDPT3 [38]; $N = 2$ is the minimum value that satisfies (16). Evidently, the SOS method is much more computationally expensive than our methods for high-degree relaxations. Rather surprisingly, MOSEK computes $T_N^{\rm in}$ more efficiently than $T_N^{\rm out}$ at large N, despite the latter being nominally cheaper. This is not the case for SDPT3.

On the other hand, Fig. 3 shows that while $T_N^{\rm sos}$ seems to converge to $T_N^{\rm out}$ as N increases, the inner approximation sets $T_N^{\rm in}$

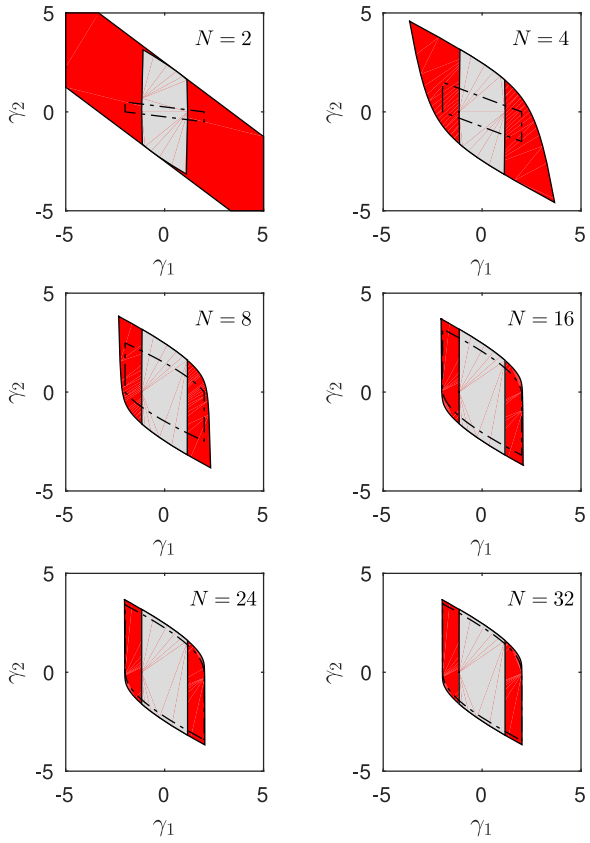


Fig. 3. (Color online) Inner and outer approximations of the feasible set of (54): T_N^{in} (black solid boundary, gray interior), T_N^{out} (black solid boundary, red interior), and T_N^{sos} (dotted–dashed black boundary, no shading).

computed with the method of Section IV do not: our estimates in Lemmas 4 and 5 and the SOS relaxation of the polynomial matrix inequality (42c) introduce conservativeness.

Yet, there are parts where the boundaries of T_N^{out} and T_N^{in} almost coincide even for N as low as 4, and—rather interestingly—these correspond to those regions where the convergence of T_N^{sos} to T_N^{out} is the slowest. The figures suggest that the inner approximation sets T_N^{in} are only overconstrained in the γ_1 direction. This is because γ_2 appears only in the term $\int_{-1}^{1} 2\gamma_2 uv \, \mathrm{d}x$, to which we apply the estimates in Lemma 5 when computing the inner SDP relaxation. According to the decay rates stated in the Lemma, however, these estimates become negligible at large N. On the contrary, γ_1 appears in the term $\int_{-1}^{1} \gamma_1 x^2 \partial u \, \partial v \, \mathrm{d}x$, to which the estimates in part 2) of Lemma 4 must be applied. Despite our efforts to tune the auxiliary matrices in (29), the magnitude of such estimates does not decay compared to other terms, limiting the range of feasible values of γ_1 in practice. This issue should be addressed in future work, and should be taken into account when trying to formulate rigorous statements on the feasibility and convergence of our inner SDP relaxations.

VII. Scalability

It may be checked that when $\boldsymbol{w}(x) \in \mathbb{R}^q$ is subject to p independent BCs, the degree of the polynomials in the matrix $\boldsymbol{F}(x;\boldsymbol{\gamma})$ is at most d_F, and $\boldsymbol{\gamma} \in \mathbb{R}^s$, the outer SDP relaxation for a quadratic inequality of the form (46) with N Legendre coefficients results in an SDP with an LMI of dimension $q(N+1)-p$ with s variables. Instead, the inner SDP relaxation has: an LMI of dimension $2|\boldsymbol{l}| + q(N + |\boldsymbol{k}|_\infty + d_F + 2) - p$, where $|\boldsymbol{k}|_\infty := \max_{i\in\{1,\ldots,q\}} k_i$; a $q \times q$ matrix SOS constraint of degree $\deg \boldsymbol{S}(x;\boldsymbol{\gamma})$, where $\boldsymbol{S}(x;\boldsymbol{\gamma})$ is defined as in Section IV-C; at most $q(q+1)/2$ auxiliary LMIs of size $4\,d_F$ from Lemma 4; at most $(d_F+1)(2\,q+|\boldsymbol{k}|+1)|\boldsymbol{k}|$ linear inequalities to lift the absolute values introduced by Lemma 5; at most $s + q(q+1)(2\,d_F^2 + d_F + 2)/2 + (d_F+1)(2\,q+|\boldsymbol{k}|+1)|\boldsymbol{k}|/2$ variables. Since currently only small- to medium-size SDPs can be solved in practice, one might therefore expect that although our techniques are cheaper than the SOS method of [3]—as highlighted by our numerical examples—they can only be implemented when q, d_F, s, and $|\boldsymbol{k}|_\infty$ are sufficiently small.

The development of solvers for large-scale SDPs is an active research area, and new tools are being developed that should facilitate solving problems at larger scales; see, for example, the solvers SCS [39] and CDCS [40], [41].

Moreover, the poor scalability of SDPs may not be too severe an issue for many problems of practical interest. In fact, the number of constraints in the inner SDP relaxation can be considerably smaller than the worst-case count presented above. To see this, note that Lemma 4 introduces auxiliary LMIs and variables only for the (upper triangular, by symmetry) entries of $\boldsymbol{S}(x;\boldsymbol{\gamma})$ that depend on x; for example, only one auxiliary LMI is needed for inequality (54). In addition, the size of the auxiliary LMI associated with the entry $\boldsymbol{S}_{ij}$ can be reduced to $4 \times \deg \boldsymbol{S}_{ij}$, yielding considerable savings if $\deg \boldsymbol{S}(x;\boldsymbol{\gamma}) \ll d_F$. In the extreme case $\deg \boldsymbol{S}(x;\boldsymbol{\gamma}) = 0$, i.e., the matrix $\boldsymbol{S}(x;\boldsymbol{\gamma})$ is independent of x, there are no auxiliary variables and LMIs from Lemma 4, and moreover the $q \times q$ matrix SOS constraint becomes a $q \times q$ LMI. This situation is common when energy-Lyapunov-function methods are applied to turbulent fluid flows [11], [13], [14], so our techniques are particularly suited to tackle problems in this field—as proven by the results of Section VI-A and of [26].

Finally, we also note that a moderate Legendre truncation parameter N, and hence a medium-size SDP relaxation, often suffices to obtain accurate bounds on the objective function, as suggested by all our examples. Roughly speaking, to obtain a good bound on the optimal value one should choose N such that the minimizer $\boldsymbol{w}^\star$ of $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ at the optimal point $\boldsymbol{\gamma} = \boldsymbol{\gamma}^\star$ is approximated sufficiently well by a polynomial of degree N (here we assume that the minimizer $\boldsymbol{w}^\star$ exists for simplicity). Since $\boldsymbol{w}^\star$ is typically a "well-behaved" function (the highest order derivatives of highly oscillatory test functions $\boldsymbol{w}$ would give large contribution to $\mathcal{F}_{\boldsymbol{\gamma}^\star}\{\boldsymbol{w}\}$, making highly oscillatory minimizers unlikely), this can be done with moderate N.

VIII. Conclusion

In this paper, we have developed a new method to optimize a linear cost function subject to homogeneous quadratic integral inequality constraints. More precisely, we have employed Legendre series expansions and functional estimates to derive inner and outer approximations of the feasible set of an integral inequality, and have shown that upper and lower bounds for the optimal cost value can be computed efficiently using semidefinite programming. We have proven that the lower bounds obtained

with our outer approximations form a nondecreasing sequence that converges to the exact optimal cost value (if this is attained). Unfortunately, similar statements do not generally extend to our inner approximations.

Although the steps leading to our SDP relaxations are rather technical, they are amenable to numerical implementation. To aid the formulation and solution of optimization problem with integral inequality constraints in practice, we have developed the MATLAB package QUINOPT, an open-source add-on for the optimization toolbox YALMIP. Using this software, we have successfully solved nontrivial problems that arise when studying the stability of autonomous systems of PDEs.

We have demonstrated that our methods work well in practice, even though they rely on typically conservative estimates to formulate numerically tractable constraints. It is in the interest of future work to formalize these observations, and determine conditions that ensure the feasibility and/or convergence of our inner SDP relaxations. The results presented in Section VI-C suggest that more stringent assumption on the properties of the integral inequality might be needed.

Looking at the applications we have in mind, i.e., the analysis of systems governed by PDEs, this work should be extended to 1) integral inequalities with explicit time dependence that arise from nonautonomous PDEs, and 2) inequalities over two or higher dimensional domains. Polynomial explicit time dependence could be dealt with by relaxing our inner/outer LMI constraints, now time-dependent, into matrix SOS conditions, although the (current) poor scalability of SOS optimization makes this strategy unlikely implementable. Multidimensional compact box domains could be analyzed by introducing Legendre expansions in each coordinate direction and adapting the ideas presented in this work, while for more general domains—including the noncompact case—other basis functions could be used. This may present hurdles in the derivation of inner approximations, because they require estimates that rely on specific properties of the basis functions. Unless sparsity and/or problem structure are exploited, multidimensional inequalities are also likely to be constrained by the current computational limitations: with n spatial dimensions and q dependent variables ($\boldsymbol{w} \in \mathbb{R}^q$), the LMI size for a simple outer approximations using polynomials of degree N will be approximately qN^n.

Finally, it is in the interest of future work to extend our methods to integral inequalities more general than the homogeneous quadratic type. We expect that our methods can be extended with little effort to complete (i.e., inhomogeneous) quadratic integral inequalities over spaces described by homogeneous BCs (inhomogeneous BCs can be "lifted" by a polynomial shift). In fact, the linear part of a complete quadratic functional can be analyzed with ideas similar to those used in Section V-A. Extensions to higher than quadratic functionals, e.g., by introducing additional slack variables to reduce them to quadratic ones, are also essential if recently developed analysis techniques based on dissipation inequalities [5] are to be successfully applied to complex nonlinear systems of PDEs of interest in physics and engineering.

Appendix A Legendre Polynomials and Legendre Series

The Legendre polynomial of degree n is defined over the interval $[-1, 1]$ as

$$\mathcal{L}_n(x) = \frac{1}{n!\,2^n}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2-1)^n.$$

The Legendre polynomials of degree $n \geq 2$ can also be constructed with the recurrence relation

$$n\mathcal{L}_n(x) = (2n-1)\,x\,\mathcal{L}_{n-1}(x) - (n-1)\mathcal{L}_{n-2}(x) \tag{55}$$

with $\mathcal{L}_0(x) = 1$ and $\mathcal{L}_1(x) = x$. Equation (55) can be used to show that $\mathcal{L}_n(\pm 1) = (\pm 1)^n$.

The Legendre polynomials satisfy a number of other recurrence relations. In this work, we will use the fact that

$$(2n+1)\mathcal{L}_n(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\mathcal{L}_{n+1}(x) - \mathcal{L}_{n-1}(x)\right], \quad n \geq 1 \tag{56}$$

see e.g., [32, Chapter 7, Problem 7.8]. Moreover, $\|\mathcal{L}_n\|_\infty \leq 1$ for all $n \geq 0$.

The Legendre polynomials also form a complete orthogonal basis for the Lebesgue space $L^2(-1,1)$ [31], and satisfy the orthogonality condition

$$\int_{-1}^{1} \mathcal{L}_n\,\mathcal{L}_m\,\mathrm{d}x = \frac{2\delta_{mn}}{2n+1} \tag{57}$$

where δ_{mn} is the usual Kronecker delta. This means that any square-integrable function u can be expanded with a convergent series (in the L^2 norm sense)

$$u(x) = \sum_{n=0}^{\infty} \hat{u}_n \mathcal{L}_n(x), \quad \hat{u}_n = \int_{-1}^{1} u\,\mathcal{L}_n\,\mathrm{d}x \tag{58}$$

where the $\hat{u}_n$'s are known as Legendre coefficients. From (57) it follows that

$$\|u\|_2^2 = \int_{-1}^{1} |u|^2\,\mathrm{d}x = \sum_{n=0}^{\infty} \frac{2|\hat{u}_n|^2}{2n+1}. \tag{59}$$

Finally, if u is continuously differentiable on $[-1,1]$ its Legendre series expansion converges uniformly. In fact, u is Lipschitz on $[-1,1]$ because, by Taylor's theorem, for any $x, y \in [-1,1]$ there exists a point z between x and y, such that $|u(y)-u(x)| = |\partial u(z)|\,|x-y| \leq C\,|x-y|$. Here, C is a generic positive constant whose existence is guaranteed by the continuity of ∂u in $[-1,1]$. Uniform convergence follows from [30, Theorem XI and subsequent comments].

Appendix B Proofs

A. Proof of Theorem 1

Define the norm $\|\boldsymbol{w}\|_k^2 := \int_{-1}^{1} (\mathcal{D}^k\boldsymbol{w})^T\mathcal{D}^k\boldsymbol{w}\,\mathrm{d}x$, consider the functional $\mathcal{H}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}/\|\boldsymbol{w}\|_k^2$ and let

$$t(\boldsymbol{\gamma}) := \inf_{\boldsymbol{w}\in H\setminus\{\boldsymbol{0}\}} \mathcal{H}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}, \quad t_N(\boldsymbol{\gamma}) := \inf_{\boldsymbol{w}\in S_N\setminus\{\boldsymbol{0}\}} \mathcal{H}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}.$$

(We need not assume that these infima are achieved.) It is not too difficult to show that the sets T and T_N^{out} are described by the inequalities $t(\boldsymbol{\gamma}) \geq 0$ and $t_N(\boldsymbol{\gamma}) \geq 0$, respectively. To prove Theorem 1, we rely on the following result.

Lemma 6: Suppose $\boldsymbol{\gamma} \notin T$, i.e., there exists $\varepsilon_{\boldsymbol{\gamma}} > 0$, such that $t(\boldsymbol{\gamma}) \leq -2\varepsilon_{\boldsymbol{\gamma}}$. Then, there exists an integer $N_{\boldsymbol{\gamma}}$, such that $t_N(\boldsymbol{\gamma}) \leq -\varepsilon_{\boldsymbol{\gamma}}$ for all $N \geq N_{\boldsymbol{\gamma}}$.

Proof: Let $\{\boldsymbol{w}_n\}_{n\geq 0}, \boldsymbol{w}_n \in H, \boldsymbol{w}_n \neq \boldsymbol{0}$ be a minimizing sequence, i.e., such that $\lim_{n\to\infty} \mathcal{H}_{\boldsymbol{\gamma}}\{\boldsymbol{w}_n\} = t(\boldsymbol{\gamma})$, and for each n define $\mu_n := \|\boldsymbol{w}_n\|_k^2/(n+1)$. Note that $F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)\right)$—the integrand of $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$—and the product $(\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x))^T\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)$

are continuous with respect to all entries of the vector $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)$ at each fixed $x \in [-1,1]$. Using $|\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} - \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}_n\}| \leq 2\left\|F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\right) - F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}_n\right)\right\|_\infty$ and a similar inequality for $\left|\|\boldsymbol{w}\|_k^2 - \|\boldsymbol{w}_n\|_k^2\right|$ it is then not difficult to show that there exists $\delta_n > 0$, such that

$$\max_{0\leq\alpha\leq k} \|\partial^\alpha \boldsymbol{w} - \partial^\alpha \boldsymbol{w}_n\|_\infty \leq \delta_n \tag{60}$$

implies

$$|\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} - \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}_n\}| \leq \mu_n, \quad \left|\|\boldsymbol{w}\|_k^2 - \|\boldsymbol{w}_n\|_k^2\right| \leq \mu_n. \tag{61}$$

Since the Weierstrass approximation theorem can be extended to linear subspaces of continuously differentiable functions with prescribed BCs (this follows e.g. from [42, Proposition 2]), there exists a polynomial $\boldsymbol{P}_n \in H$ of degree d_n, $\boldsymbol{P}_n \neq \boldsymbol{0}$, that satisfies (60). Without loss of generality, we may assume that $d_n < d_{n+1}$. From (61), we see that $\|\boldsymbol{P}_n\|_k^2 \geq \|\boldsymbol{w}_n\|_k^2 - \mu_n = \|\boldsymbol{w}_n\|_k^2 \times n/(n+1)$, and since $\boldsymbol{P}_n \in S_N$ for all $N \in \{d_n, \ldots, d_{n+1}-1\}$ we can write

$$\begin{aligned} t(\boldsymbol{\gamma}) \leq t_N(\boldsymbol{\gamma}) &\leq \mathcal{H}_{\boldsymbol{\gamma}}\{\boldsymbol{P}_n\} \\ &\leq \frac{|\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{P}_n\} - \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}_n\}|}{\min\{\|\boldsymbol{P}_n\|_k^2\,,\, \|\boldsymbol{w}_n\|_k^2\}} + \mathcal{H}_{\boldsymbol{\gamma}}\{\boldsymbol{w}_n\} \\ &\leq \frac{1}{n} + \mathcal{H}_{\boldsymbol{\gamma}}\{\boldsymbol{w}_n\}. \end{aligned}$$

The last expression tends to $t(\boldsymbol{\gamma})$ as n (hence, N) tends to infinity, so $t_N(\boldsymbol{\gamma}) - t(\boldsymbol{\gamma}) \downarrow 0$ as $N \to \infty$. In particular, there exists an integer $N_{\boldsymbol{\gamma}}$, such that $t_N(\boldsymbol{\gamma}) - t(\boldsymbol{\gamma}) \leq \varepsilon_{\boldsymbol{\gamma}}$ for all $N \geq N_{\boldsymbol{\gamma}}$. Upon rearranging and recalling that $t(\boldsymbol{\gamma}) \leq -2\varepsilon_{\boldsymbol{\gamma}}$ we conclude that, for all $N \geq N_{\boldsymbol{\gamma}}$

$$t_N(\boldsymbol{\gamma}) \leq \varepsilon_{\boldsymbol{\gamma}} + t(\boldsymbol{\gamma}) \leq \varepsilon_{\boldsymbol{\gamma}} - 2\varepsilon_{\boldsymbol{\gamma}} = -\varepsilon_{\boldsymbol{\gamma}} < 0.$$

■

Let us now prove Theorem 1. The sequence of optimal values $\{p_N^\star\}_{N\geq 0}$ is nondecreasing since $T_{N+1}^{\text{out}} \subset T_N^{\text{out}}$. To prove convergence when (6) achieves its optimal value, let us assume that its feasible set T is *bounded*; if not, one can formulate an equivalent problem (meaning that $\boldsymbol{\gamma}^\star$ is still an optimal solution) with bounded feasible set by adding the constraint $\|\boldsymbol{\gamma}\| \leq r$ for a sufficiently large $r > 0$. For any $\varepsilon > 0$ (different from that used in Lemma 6), the set

$$K := \{\boldsymbol{\gamma} : \varepsilon \leq \text{dist}(\boldsymbol{\gamma}, T) \leq 2\,\varepsilon\}$$

where $\text{dist}(\boldsymbol{\gamma}, T) = \min_{\boldsymbol{\eta}\in T} \|\boldsymbol{\eta} - \boldsymbol{\gamma}\|$ is the usual euclidean distance of $\boldsymbol{\gamma}$ from T, is compact, and only contains points that are infeasible for (6). By Lemma 6, for each $\boldsymbol{\gamma} \in K$, there exists an integer $N_{\boldsymbol{\gamma}}$, such that $t_N(\boldsymbol{\gamma}) < 0$, i.e., $\boldsymbol{\gamma}$ is infeasible for (13), for all $N \geq N_{\boldsymbol{\gamma}}$. The compactness of K and the continuity of $t_N(\boldsymbol{\gamma})$—the proof of this fact is not difficult and is left to the reader—imply the existence of an integer $N_0 = N_0(\varepsilon)$ and a finite number of balls $B(\boldsymbol{\gamma}_i, \delta_i)$ with center $\boldsymbol{\gamma}_i$ and radius δ_i which cover K such that $t_N(\boldsymbol{\gamma}) < 0$ in each ball for all $N \geq N_0$. Consequently, all points in K are infeasible for the outer SDP relaxation (13) when $N \geq N_0$. Since the feasible set T_N^{out} of the outer SDP relaxation must be convex, we conclude that it must be contained within an ε-neighborhood of T for all $N \geq N_0$, i.e.,

$$\forall N \geq N_0, \quad \max_{\boldsymbol{\gamma}\in T_N^{\text{out}}} \text{dist}(\boldsymbol{\gamma}, T) < \varepsilon.$$

In particular, T_N^{out} is bounded, and there exists a point $\boldsymbol{\gamma}_N^\star$ with $\boldsymbol{c}^T\boldsymbol{\gamma}_N^\star = p_N^\star$ whose projection onto T, denoted $\mathcal{P}_T(\boldsymbol{\gamma}_N^\star)$, satisfies $\|\mathcal{P}_T(\boldsymbol{\gamma}_N^\star) - \boldsymbol{\gamma}_N^\star\| < \varepsilon$. Then, for all $N \geq N_0$

$$\begin{aligned} p^\star - \|\boldsymbol{c}\|\,\varepsilon &\leq \boldsymbol{c}^T\mathcal{P}_T(\boldsymbol{\gamma}_N^\star) - \|\boldsymbol{c}\|\,\varepsilon \\ &= \boldsymbol{c}^T\left[\mathcal{P}_T(\boldsymbol{\gamma}_N^\star) - \boldsymbol{\gamma}_N^\star\right] + p_N^\star - \|\boldsymbol{c}\|\,\varepsilon \\ &\leq \|\boldsymbol{c}\|\,\|\mathcal{P}_T(\boldsymbol{\gamma}_N^\star) - \boldsymbol{\gamma}_N^\star\| + p_N^\star - \|\boldsymbol{c}\|\,\varepsilon \\ &< p_N^\star. \end{aligned}$$

Since $p_N^\star \leq p^\star$, we conclude that $p^\star - \|\boldsymbol{c}\|\,\varepsilon \leq \lim_{N\to\infty} p_N^\star \leq p^\star$ for any ε, and the proof is concluded by letting $\varepsilon \to 0$.

B. Proof of Lemma 1

The statement is trivial when $\alpha = k$. Moreover, since $u \in C^m([-1,1])$ and $k \leq m-1$, the Legendre expansions of all derivatives $\partial^\alpha u$, $\alpha \in \{0,\ldots,k\}$ converge uniformly, cf. Appendix A. Consequently, we can use the fundamental theorem of calculus for each $\alpha \leq k-1$ to write

$$\begin{aligned} (\partial^\alpha u)(x) &= \partial^\alpha u|_{-1} + \int_{-1}^x \partial^{\alpha+1}u(t)\,\mathrm{d}t \\ &= \partial^\alpha u|_{-1} + \sum_{n\geq 0} \hat{u}_n^{\alpha+1} \int_{-1}^x \mathcal{L}_n(t)\,\mathrm{d}t. \end{aligned} \tag{62}$$

The last expression can be integrated recalling that $\mathcal{L}_0(x) = 1$, $\mathcal{L}_1(x) = x$, $\mathcal{L}_n(\pm 1) = (\pm 1)^n$ and using the recurrence relation (56). We can then rewrite (62) as

$$\partial^\alpha u = \partial^\alpha u|_{-1} + [\mathcal{L}_1 + \mathcal{L}_0]\,\hat{u}_0^{\alpha+1} + \sum_{n\geq 1}[\mathcal{L}_{n+1} - \mathcal{L}_{n-1}]\,\frac{\hat{u}_n^{\alpha+1}}{2n+1}.$$

Rearranging the series and comparing coefficients with the Legendre expansion of $\partial^\alpha u$ gives the relations

$$\hat{u}_0^\alpha = \partial^\alpha u|_{-1} + \hat{u}_0^{\alpha+1} - \frac{1}{3}\hat{u}_1^{\alpha+1}, \tag{63a}$$

$$\hat{u}_n^\alpha = \frac{\hat{u}_{n-1}^{\alpha+1}}{2n-1} - \frac{\hat{u}_{n+1}^{\alpha+1}}{2n+3}, \qquad n \geq 1. \tag{63b}$$

We can then find matrices $\boldsymbol{C}^\alpha$ and $\boldsymbol{E}^\alpha$, such that

$$\hat{\boldsymbol{u}}_{[r,s]}^\alpha = \boldsymbol{E}^\alpha\mathcal{D}^{k-1}u|_{-1} + \boldsymbol{C}^\alpha\hat{\boldsymbol{u}}_{[r-1,s+1]}^{\alpha+1}. \tag{64}$$

Here and in the following, it should be understood that negative indices should be replaced by 0. Before proceeding, note that strictly speaking the matrices $\boldsymbol{C}^\alpha$ and $\boldsymbol{E}^\alpha$ depend on r and s, but we do not write this explicitly to ease the notation. In particular, (63) implies that $\boldsymbol{E}^\alpha = \boldsymbol{0}$ if $r \geq 1$.

Expressions similar to (64) can be built for all vectors $\hat{\boldsymbol{u}}_{[r-i,s+i]}^{\alpha+i}$, $i \in \{0,\ldots,k-\alpha-1\}$. After some algebra, it is therefore possible to write

$$\hat{\boldsymbol{u}}_{[r,s]}^\alpha = \boldsymbol{B}_{[r,s]}^\alpha\mathcal{D}^{k-1}u|_{-1} + \left(\prod_{i=0}^{k-\alpha-1}\boldsymbol{C}^{\alpha+i}\right)\hat{\boldsymbol{u}}_{[r-k+\alpha,s+k-\alpha]}^k \tag{65}$$

where

$$\boldsymbol{B}^{\alpha}_{[r,s]} := \boldsymbol{E}^{\alpha} + \boldsymbol{C}^{\alpha}\boldsymbol{E}^{\alpha+1} + \cdots + \left(\prod_{i=0}^{k-\alpha-2} \boldsymbol{C}^{\alpha+i}\right)\boldsymbol{E}^{k-1}.$$

Note that, in light of (63), all matrices $\boldsymbol{E}^{\alpha+i}$, $i \in \{0, \ldots, k-\alpha-1\}$, are zero if $r \geq k-\alpha$. Since we have assumed that $s+k-\alpha \leq M$, the last term in (65) can be rewritten in terms of $\hat{\boldsymbol{u}}^k_{[0,M]}$ (recall that $r-k+\alpha$ is replaced by 0 if it is negative). The proof is concluded by defining $\boldsymbol{D}^{\alpha}_{[r,s]}$ as in (66) shown at the bottom of this page, where the size of the zero matrices is indicated by subscripts.

C. Proof of Lemma 2

Recalling the definition of $\mathcal{B}^{k-1}u$, we only need to show that $\mathcal{D}^{k-1}u|_1$ can be expressed as linear combination of the entries of $\check{\boldsymbol{u}}_M$. Applying the fundamental theorem of calculus as in Appendix B-B, it may be shown that $\partial^{\alpha}u|_1 = \partial^{\alpha}u|_{-1} + 2\hat{u}_0^{\alpha+1}$ for any $\alpha \in \{0, \ldots, k-1\}$. By Lemma 1, $\partial^{\alpha}u|_1$ can then be written as a linear combination of the entries of $\check{\boldsymbol{u}}_M$. Repeating this argument for all $\alpha \in \{0, \ldots, k-1\}$, we conclude the same for all entries of $\mathcal{D}^{k-1}u|_1$, proving the existence of $\boldsymbol{G}_M$.

D. Proof of Lemma 4

1) Recall (15) and expand

$$\mathcal{Q}^{\alpha\beta}_{uv} = \sum_{m=0}^{N_\alpha}\sum_{n=N_\beta+1}^{+\infty} \hat{u}^{\alpha}_m \hat{v}^{\beta}_n \int_{-1}^{1} f\,\mathcal{L}_m\,\mathcal{L}_n\,\mathrm{d}x + \sum_{m=N_\alpha+1}^{\infty}\sum_{n=0}^{N_\beta} \hat{u}^{\alpha}_m \hat{v}^{\beta}_n \int_{-1}^{1} f\,\mathcal{L}_m\,\mathcal{L}_n\,\mathrm{d}x$$

where $N_\alpha = N+\alpha$ and $N_\beta = N+\beta$. Since f is a polynomial of degree at most d_F, the product $f\mathcal{L}_m$ is a polynomial of degree at most $m+d_F$, so it is orthogonal to any Legendre polynomial $\mathcal{L}_n$ with $n > m+d_F$. In particular, it may be shown [43] that the integral $\int_{-1}^{1} f\mathcal{L}_n\mathcal{L}_m\,\mathrm{d}x$ vanishes if $|m-n| > d_F$. Using the shorthand notation $\overline{n} = n+1-d_F$, we can write (67) shown at the bottom of this page. Note that we have assumed that α, β, and d_F are such that $1-d_F \leq \alpha-\beta \leq d_F-1$, so that the vectors in (67) are well defined. If the left (resp. right) inequality is not satisfied, then the first (resp. second) term in (67) vanishes. Since $N_\alpha + d_F \leq M+\beta-k$ and $N_\beta + d_F \leq M+\alpha-k$, we can apply Lemma 1, and our assumption that $N \geq d_F+k-1$ guarantees that $\overline{N_\alpha} \geq k-\beta$ and $\overline{N_\beta} \geq k-\alpha$, so there is no dependence on the boundary values. Consequently, we can find a matrix $\boldsymbol{Q}(\boldsymbol{\gamma})$, such that

$$\mathcal{Q}^{\alpha\beta}_{uv} = \left(\hat{\boldsymbol{u}}^k_{[0,M]}\right)^T \boldsymbol{Q}(\boldsymbol{\gamma})\,\hat{\boldsymbol{v}}^k_{[0,M]}.$$

The matrix $\boldsymbol{Q}^{\alpha\beta}_{uv}$ is found using (26) after taking the symmetric part of the right-hand side of the last expression.

2) Let $\boldsymbol{\nu} = \left[\,\hat{u}^k_{M+1}, \ldots, \hat{u}^k_{M+d_F}, \hat{v}^k_{M+1}, \ldots, \hat{v}^k_{M+d_F}\,\right]^T$. After replacing N_α and N_β with M in (67), it may be verified using (26) that $\mathcal{Q}^{kk}_{uv} = 2\,{\boldsymbol{\psi}_M}^T {\boldsymbol{L}_{\overline{M}}}^T \boldsymbol{Y}\boldsymbol{\nu}$. By (28)

$$0 \leq \begin{bmatrix}\boldsymbol{L}_{\overline{M}}\boldsymbol{\psi}_M \\ \boldsymbol{\nu}\end{bmatrix}^T \begin{bmatrix}\boldsymbol{Q}^{kk}_{uv} & \boldsymbol{Y} \\ \boldsymbol{Y}^T & \boldsymbol{\Sigma}^{kk}_{uv}\otimes\boldsymbol{\Delta}\end{bmatrix}\begin{bmatrix}\boldsymbol{L}_{\overline{M}}\boldsymbol{\psi}_M \\ \boldsymbol{\nu}\end{bmatrix} = {\boldsymbol{\psi}_M}^T \times \left({\boldsymbol{L}_{\overline{M}}}^T\boldsymbol{Q}^{kk}_{uv}\boldsymbol{L}_{\overline{M}}\right)\boldsymbol{\psi}_M + \boldsymbol{\nu}^T\left(\boldsymbol{\Sigma}^{kk}_{uv}\otimes\boldsymbol{\Delta}\right)\boldsymbol{\nu} + \mathcal{Q}^{kk}_{uv}.$$

Now, $\boldsymbol{\Sigma}^{kk}_{uv}$ is a diagonal matrix by assumption. Recalling the definition of $\boldsymbol{\Delta}$, $\boldsymbol{\nu}$, and rearranging we arrive at (68) shown at the bottom of the this page. Recognizing from (59) that the two sums in (68) can be bounded by $\left\|U^k_M\right\|_2^2$ and $\left\|V^K_M\right\|_2^2$, we obtain (29).

E. Proof of Lemma 5

We start by determining an upper bound on $\|U^{\alpha}_{N_\alpha}\|_2^2$ in terms of the vector $\hat{\boldsymbol{u}}^k_{[0,M]}$ and $\|U^k_M\|_2^2$ (similar bounds can be found for $V^{\beta}_{N_\beta}$). Recalling (15), (16), and (19), we can write

$$\frac{1}{2}\left\|U^{\alpha}_{N_\alpha}\right\|_2^2 = \sum_{n=N_\alpha+1}^{M-k+\alpha}\frac{(\hat{u}^{\alpha}_n)^2}{2n+1} + \sum_{n=M-k+\alpha+1}^{+\infty}\frac{(\hat{u}^{\alpha}_n)^2}{2n+1} = \left(\hat{\boldsymbol{u}}^k_{[0,M]}\right)^T \boldsymbol{H}_{\alpha}\,\hat{\boldsymbol{u}}^k_{[0,M]} + \sum_{n=M-k+\alpha+1}^{+\infty}\frac{(\hat{u}^{\alpha}_n)^2}{2n+1} \quad (69)$$

where the matrix $\boldsymbol{H}_{\alpha}$ can be obtained from Lemma 1. In particular, we note that (63b) is applied $k-a$ times to $(\hat{u}^{\alpha}_n)^2$ to compute $\boldsymbol{H}_{\alpha}$, and since $n > N_\alpha \geq N$ it may be verified that $\|\boldsymbol{H}_{\alpha}\|_F \sim N^{-2(k-\alpha)-1}$.

When $\alpha = k$, the last term in (69) is $\left\|U^k_M\right\|_2^2/2$, so

$$\frac{1}{2}\left\|U^k_{N_k}\right\|_2^2 = \left(\hat{\boldsymbol{u}}^k_{[0,M]}\right)^T \boldsymbol{H}_k\,\hat{\boldsymbol{u}}^k_{[0,M]} + \frac{1}{2}\left\|U^k_M\right\|_2^2. \quad (70)$$

$$\boldsymbol{D}^{\alpha}_{[r,s]} := \left[\boldsymbol{0}_{(s-r+1)\times(r-k+\alpha)}, \quad \prod_{i=0}^{k-\alpha-1}\boldsymbol{C}^{\alpha+i}, \quad \boldsymbol{0}_{(s-r+1)\times(M-s-k+\alpha)}\right]. \quad (66)$$

$$\mathcal{Q}^{\alpha\beta}_{uv} = \begin{bmatrix}\hat{u}^{\alpha}_{\overline{N_\beta}} \\ \vdots \\ \hat{u}^{\alpha}_{N_\alpha}\end{bmatrix}^T \boldsymbol{\Phi}^{[N_\beta+1,N_\alpha+d_F]}_{[\overline{N_\beta},N_\alpha]} \begin{bmatrix}\hat{v}^{\beta}_{N_\beta+1} \\ \vdots \\ \hat{v}^{\beta}_{N_\alpha+d_F}\end{bmatrix} + \begin{bmatrix}\hat{v}^{\beta}_{\overline{N_\alpha}} \\ \vdots \\ \hat{v}^{\beta}_{N_\beta}\end{bmatrix}^T \boldsymbol{\Phi}^{[N_\alpha+1,N_\beta+d_F]}_{[\overline{N_\alpha},N_\beta]} \begin{bmatrix}\hat{u}^{\alpha}_{N_\alpha+1} \\ \vdots \\ \hat{u}^{\alpha}_{N_\beta+d_F}\end{bmatrix}. \quad (67)$$

$$\mathcal{Q}^{kk}_{uv} \geq -{\boldsymbol{\psi}_M}^T\left({\boldsymbol{L}_{\overline{M}}}^T\boldsymbol{Q}^{kk}_{uv}\boldsymbol{L}_{\overline{M}}\right)\boldsymbol{\psi}_M - (\boldsymbol{\Sigma}^{kk}_{uv})_{11}\sum_{n=M+1}^{M+d_F}\frac{2|\hat{u}^k_n|^2}{2n+1} - (\boldsymbol{\Sigma}^{kk}_{uv})_{22}\sum_{n=M+1}^{M+d_F}\frac{2|\hat{v}^k_n|^2}{2n+1}. \quad (68)$$

$$\sum_{n=M-k+\alpha+1}^{+\infty} \frac{(\hat{u}_n^\alpha)^2}{2n+1} \leq \sum_{n=M-k+\alpha+1}^{+\infty} \frac{2}{2n+1}\left[\frac{|\hat{u}_{n-1}^{\alpha+1}|^2}{(2n-1)^2} + \frac{|\hat{u}_{n+1}^{\alpha+1}|^2}{(2n+3)^2}\right]$$
$$= \sum_{n=M-k+\alpha}^{M-k+\alpha+1} \frac{2|\hat{u}_n^{\alpha+1}|^2}{(2n+3)(2n+1)^2} + \sum_{n=M-k+\alpha+2}^{\infty} \frac{4|\hat{u}_n^{\alpha+1}|^2}{(2n-1)(2n+1)(2n+3)}$$
$$\leq \sum_{n=M-k+\alpha}^{M-k+\alpha+1} \frac{2|\hat{u}_n^{\alpha+1}|^2}{(2n+3)(2n+1)^2} + \omega_{\alpha+1} \sum_{n=M-k+\alpha+2}^{+\infty} \frac{|\hat{u}_n^{\alpha+1}|^2}{2n+1}. \tag{72}$$

When $\alpha \leq k-1$, instead, we define

$$\omega_\eta = \frac{4}{[2(M-k+\eta)+1][2(M-k+\eta)+5]} \tag{71}$$

for $\eta \in \{0, \ldots, k-1\}$ and use (63), the elementary inequality $(a-b)^2 \leq 2(a^2+b^2)$, and appropriate changes of indices to show (72) shown at the top of this page. Applying Lemma 1 to the first term on the right-hand side of (72) and substituting back into (69), we can construct a matrix $\boldsymbol{T}_\alpha$, such that

$$\frac{1}{2}\left\|U_{N_\alpha}^\alpha\right\|_2^2 \leq \left(\hat{\boldsymbol{u}}_{[0,M]}^k\right)^T \boldsymbol{T}_\alpha\, \hat{\boldsymbol{u}}_{[0,M]}^k + \omega_{\alpha+1} \sum_{n=M-k+\alpha+2}^{+\infty} \frac{|\hat{u}_n^{\alpha+1}|^2}{2n+1}. \tag{73}$$

As for $\boldsymbol{H}_\alpha$, it may be verified that $\|\boldsymbol{T}_\alpha\|_F \sim N^{-2(k-\alpha)-1}$.

Similar estimates can be carried out for the infinite sum on the right-hand side of (73). By recursion, we can eventually construct a matrix $\boldsymbol{Z}_\alpha$ and a constant λ_α, such that

$$\frac{1}{2}\left\|U_{N_\alpha}^\alpha\right\|_2^2 \leq \left(\hat{\boldsymbol{u}}_{[0,M]}^k\right)^T \boldsymbol{Z}_\alpha\, \hat{\boldsymbol{u}}_{[0,M]}^k + \lambda_\alpha \left\|U_M^k\right\|_2^2. \tag{74}$$

Note that $\|\boldsymbol{Z}_\alpha\|_F \sim N^{-2(k-\alpha)-1}$, while $\lambda_\alpha \sim N^{-2(k-\alpha)}$ since every recursion step introduces a factor of N^{-2} according to (71). Moreover, the right-hand side of (74) has the same form as (70), so for the rest of this section we will not distinguish the cases $\alpha \leq k-1$ and $\alpha = k$.

The estimate (74) can be used in conjunction with Young's inequality and (26) to show that for any $\varepsilon > 0$ we can bound

$$|\mathcal{R}_{uv}^{\alpha\beta}| \leq \|f\|_\infty \boldsymbol{\psi}_M{}^T \left(\boldsymbol{L}_0{}^T \begin{bmatrix} \varepsilon \boldsymbol{Z}_\alpha & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\varepsilon}\boldsymbol{Z}_\beta \end{bmatrix} \boldsymbol{L}_0\right) \boldsymbol{\psi}_M{}^T + \|f\|_\infty \left(\varepsilon\lambda_\alpha \|U_k\|_2^2 + \frac{\lambda_\beta}{\varepsilon}\|V_k\|_2^2\right).$$

We now set $\varepsilon = (N+1)^{\beta-\alpha}$, so $\varepsilon\lambda_\alpha \sim \varepsilon^{-1}\lambda_\beta \sim N^{\alpha+\beta-2k}$ and $\|\varepsilon \boldsymbol{Z}_\alpha\|_F \sim \left\|\varepsilon^{-1}\boldsymbol{Z}_\beta\right\|_F \sim N^{\alpha+\beta-2k-1}$, and let

$$\boldsymbol{R}_{uv}^{\alpha\beta} := \boldsymbol{L}_0{}^T \begin{bmatrix} \varepsilon \boldsymbol{Z}_\alpha & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\varepsilon}\boldsymbol{Z}_\beta \end{bmatrix} \boldsymbol{L}_0, \quad \boldsymbol{\Sigma}_{uv}^{\alpha\beta} := \begin{bmatrix} \varepsilon\lambda_\alpha & 0 \\ 0 & \varepsilon^{-1}\lambda_\beta \end{bmatrix}.$$

Recalling that $|\mathcal{L}_n(x)| \leq 1$ for all $n \geq 0$ [30], (30) then follows from the estimate

$$\|f\|_\infty = \sup_{x\in[-1,1]} \left|\sum_{n=0}^p \hat{f}_n(\gamma)\mathcal{L}_n(x)\right| \leq \sum_{n=0}^p \left|\hat{f}_n(\gamma)\right| = \left\|\hat{\boldsymbol{f}}(\gamma)\right\|_1.$$

REFERENCES

[1] B. Straughan, *The Energy Method, Stability, and Nonlinear Convection*, (Applied Mathematical Sciences), vol. 91, 2nd ed. New York, NY, USA: Springer-Verlag, 2004.

[2] G. Valmorbida, M. Ahmadi, and A. Papachristodoulou, "Semi-definite programming and functional inequalities for Distributed Parameter Systems," in *Proc. 53rd IEEE Conf. Decis. Control*, Los Angeles, CA, USA, 2014, pp. 4304–4309.

[3] G. Valmorbida, M. Ahmadi, and A. Papachristodoulou, "Stability analysis for a class of partial differential equations via semidefinite programming," *IEEE Trans. Automat. Control*, vol. 61, no. 6, pp. 1649–1654, Jun. 2016.

[4] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, "Input-output analysis of distributed parameter systems using convex optimization," in *Proc. IEEE 53rd Annu. Conf. Decis. Control*, 2014, pp. 4310–4315.

[5] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, "Dissipation inequalities for the analysis of a class of PDEs," *Automatica*, vol. 66, pp. 163–171, 2016.

[6] D. Huang, S. Chernyshenko, D. Lasagna, and O. Tutty, "Long-time average cost control of polynomial systems : A sum of squares approach," in *Proc. Eur. Control Conf.*, Linz, Austria, 2015, pp. 3244–3249.

[7] D. Huang and S. Chernyshenko, "Low-order state-feedback controller design for long-time average cost control of fluid flow systems: A sum-of-squares approach," in *Proc. 34th Chin. Control Conf.*, Hangzhou, China, 2015, pp. 2479–2484.

[8] D. Lasagna, D. Huang, O. R. Tutty, and S. Chernyshenko, "Controlling fluid flows with positive polynomials," in *Proc. 35th Chin. Control Conf.*, Chengdu, China, 2016, pp. 1301–1306.

[9] D. Lasagna, D. Huang, O. R. Tutty, and S. I. Chernyshenko, "A sum-of-squares approach to feedback control of laminar wake flows," *J. Fluid Mech.*, vol. 809, pp. 628–663, 2016.

[10] P. Constantin and C. R. Doering, "Variational bounds in dissipative systems," *Phys. D Nonlinear Phenom.*, vol. 82, no. 3, pp. 221–228, 1995.

[11] P. Constantin and C. R. Doering, "Variational bounds on energy dissipation in incompressible flows. II. Channel flow," *Phys. Rev. E*, vol. 51, no. 4, pp. 3192–3198, 1995.

[12] C. R. Doering and P. Constantin, "Energy dissipation in shear driven turbulence," *Phys. Rev. Lett.*, vol. 69, no. 11, pp. 1648–1651, 1992.

[13] C. R. Doering and P. Constantin, "Variational bounds on energy dissipation in incompressible flows: Shear flow," *Phys. Rev. E*, vol. 49, no. 5, pp. 4087–4099, 1994.

[14] C. R. Doering and P. Constantin, "Variational bounds on energy dissipation in incompressible flows. III. Convection," *Phys. Rev. E*, vol. 53, no. 6, pp. 5957–5981, May 1996.

[15] G. I. Hagstrom and C. R. Doering, "Bounds on surface stress-driven shear flow," *J. Nonlinear Sci.*, vol. 24, no. 1, pp. 185–199, 2014.

[16] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 1, 1st ed. New York, NY, USA: Interscience, 1953.

[17] M. Giaquinta and S. Hildebrandt, *Calculus of Variations I* (Grundlehren der mathematischen Wissenschaften), vol. 310. Berlin, Germany: Springer, 1996.

[18] C. R. Doering and J. M. Hyman, "Energy stability bounds on convective heat transport: Numerical study," *Phys. Rev. E*, vol. 55, no. 6, pp. 7775–7778, 1997.
[19] B. Wen, G. P. Chini, N. Dianati, and C. R. Doering, "Computational approaches to aspect-ratio-dependent upper bounds and heat flux in porous medium convection," *Phys. Lett. A*, vol. 377, no. 41, pp. 2931–2938, Dec. 2013.
[20] B. Wen, G. P. Chini, R. R. Kerswell, and C. R. Doering, "Time-stepping approach for solving upper-bound problems: Application to two-dimensional Rayleigh-Bénard convection," *Phys. Rev. E*, vol. 92, no. 4, 2015, Art. no. 043012.
[21] D. Bertsimas and C. Caramanis, "Bounds on linear PDEs via semidefinite optimization," *Math. Program. Ser. A*, vol. 108, no. 1, pp. 135–158, 2006.
[22] A. Papachristodoulou and M. M. Peet, "On the analysis of systems described by classes of partial differential equations," in *Proc. 45th IEEE Conf. Decis. Control*, 2006, pp. 747–752.
[23] G. Valmorbida and A. Papachristodoulou, "Introducing INTSOSTOOLS: A SOSTOOLS plug-in for integral inequalities," in *Proc. Eur. Control Conf.*, Linz, Austria, 2015, pp. 1231–1236.
[24] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, "A convex approach to hydrodynamic analysis," in *Proc. 54th IEEE Conf. Decis. Control*, Osaka, Japan, 2015, pp. 7262–7267.
[25] G. Fantuzzi and A. Wynn, "Construction of an optimal background profile for the Kuramoto–Sivashinsky equation using semidefinite programming," *Phys. Lett. A*, vol. 379, nos. 1/2, pp. 23–32, Jan. 2015.
[26] G. Fantuzzi and A. Wynn, "Optimal bounds with semidefinite programming: An application to stress driven shear flows," *Phys. Rev. E*, vol. 93, no. 4, 2016, Art. no. 043308.
[27] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. Comput. Aided Control Syst. Design*, Taipei, Taiwan, 2004, pp. 284–289.
[28] J. Löfberg, "Pre-and post-processing sum-of-squares programs in practice," *IEEE Trans. Automat. Control*, vol. 54, no. 5, pp. 1007–1011, May 2009.
[29] W. Tang, C. P. Caulfield, and W. R. Young, "Bounds on dissipation in stress-driven flow," *J. Fluid Mech.*, vol. 510, pp. 333–352, 2004.
[30] D. Jackson, *The Theory of Approximation* (American Mathematical Society Colloquium Publications), vol. 11. New York, NY, USA: Amer. Math. Soc., 1930.
[31] E. Zeidler, *Applied Functional Analysis - Applications to Mathematical Physics* (Applied Mathematical Sciences), vol. 108, 1st ed. New York, NY, USA: Springer-Verlag, 1995.
[32] R. P. Agarwal and D. O'Regan, *Ordinary and Partial Differential Equations, With Special Functions, Fourier Series and Boundary Value Problems*. New York, NY, USA: Springer-Verlag, 2009.
[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[34] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Math. Program. Ser. B*, vol. 96, no. 2, pp. 293–320, 2003.
[35] W. Tan and A. Packard, "Stability region analysis using sum of squares programming," in *Proc. Amer. Control Conf.*, Minneapolis, MN, USA, 2006, pp. 2297–2302.
[36] E. D. Andersen, B. Jensen, J. Jensen, and R. Sandvik, "MOSEK version 6," MOSEK ApS, Copenhagen, Denmark,. *MOSEK* Tech. Rep. TR-2009–3, 2009.
[37] G. Valmorbida, M. Ahmadi, and A. Papachristodoulou, "Semi-definite programming and functional inequalities for distributed parameter systems," 2014, arXiv:1403.6882.
[38] R. H. Tütüncü, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Math. Program. Ser. B*, vol. 95, no. 2, pp. 189–217, 2003.
[39] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic optimization via operator splitting and homogeneous self-dual embedding," *J. Optim. Theory Appl.*, vol. 169, no. 3, pp. 1–27, 2016.
[40] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. J. Goulart, and A. Wynn, "Fast ADMM for semidefinite programs with chordal sparsity," *American Control Conf.*, Seattle, USA, May. 24–26, 2017, doi: 10.23919/ACC.2017.7963462.
[41] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. Goulart, and A. Wynn, "Fast ADMM for homogeneous self-dual embeddings of sparse SDPs," in *Proc. 20th IFAC World Congress*, Toulouse, France, Jul. 9–14, 2017, arXiv:1611.01828v1.
[42] M. M. Peet and P.-A. Bliman, "An extension of the weierstrass theorem to linear varieties: Application to delayed systems," in *Proc. 7th IFAC Work. Time-Delay Syst.*, Nantes, France, 2007, pp. 1–4.
[43] J. Dougall, "The product of two Legendre polynomials," *Proc. Glasgow Math. Assoc.*, vol. 1, no. 3, pp. 121–125, 1953.



Giovanni Fantuzzi received the M.Eng. degree in aeronautical engineering from Imperial College London, London, U.K. in 2014. He is currently working towards the Ph.D. degree in aeronautics in the Department of Aeronautics at Imperial College London.

In 2015, he was a Fellow at the Geophysical Fluid Dynamics program at the Woods Hole Oceanographic Institution, Woods Hole, MA, USA. His research interests include fluid dynamics, semidefinite and sum-of-squares optimization, and the application of optimization techniques to partial differential equations.



Andrew Wynn received an MMath degree in mathematics at master level from the University of Oxford, in 2005 and the D.Phil. degree in pure mathematics from the University of Oxford, Oxford, U.K., in 2009.

He held Postdoctoral positions in the Department of Mathematics, University College London, and the Department of Aeronautics, Imperial College London, London, U.K., from 2009 to 2012. He is currently a Senior Lecturer in Control Engineering in the Department of Aeronautics, Imperial College London. His research interests include operator theory, distributed parameter systems, fluid dynamics and flow control.



Paul J. Goulart (M'04) received the B.Sc. and M.Sc. degrees in aeronautics and astronautics from the Massachusetts Institute of Technology, Cambridge, MA, USA, and the Ph.D. degree in control engineering from the University of Cambridge, Cambridge, U.K., in 2007.

He was selected as a Gates Scholar at the University of Cambridge. From 2007 to 2011, he was a Lecturer in control systems in the Department of Aeronautics, Imperial College London, London, U.K., and from 2011 to 2014 as a Senior Researcher in the Automatic Control Laboratory, ETH Zürich, Zürich, Switzerland. He is currently an Associate Professor in the Department of Engineering Science and a Tutorial Fellow at St. Edmund Hall, University of Oxford, Oxford, U.K. His research interests include model predictive control, robust optimization, and control of fluid flows.



Antonis Papachristodoulou received the M.A./M.Eng. degree in electrical and information sciences from the University of Cambridge, Cambridge, U.K., and the Ph.D. degree in control and dynamical systems (with a minor in aeronautics) from the California Institute of Technology, Pasadena, CA, USA.

He is currently an Associate Professor in Engineering Science with the University of Oxford, Oxford, U.K., and a Tutorial Fellow at Worcester College, Oxford. He is an EPSRC Fellow for Growth in Synthetic Biology and the Director of the EPSRC & BBSRC Centre for Doctoral Training in Synthetic Biology. His research interests include large-scale nonlinear systems analysis, sum of squares programming, synthetic and systems biology, networked systems, and flow control.

Dr. Papachristodoulou received the 2015 European Control Award for his contributions to robustness analysis and applications to networked control systems and systems biology. In the same year, he received the O. Hugo Schuck Best Paper Award.