Optimization with affine homogeneous quadratic integral inequality constraints*

Giovanni Fantuzzi† Andrew Wynn† Paul Goulart‡
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 given homogeneous quadratic integral functional, affine in the optimization variables, is non-negative over a space of functions defined by homogeneous boundary conditions. Problems of this type often arise when studying dynamical systems governed by partial differential equations. First, we derive a hierarchy of outer approximations for the feasible set of a homogeneous quadratic integral inequality in terms of linear matrix inequalities (LMIs), and show that a convergent, non-decreasing sequence of lower bounds for the optimal cost value can be computed by solving a sequence of semidefinite programs (SDPs). Second, we obtain inner approximations in terms of LMIs and sum-of-squares constraints, which enable us to formulate SDPs to compute upper bounds for the optimal cost, as well as to compute a strictly feasible point for the integral inequality. To aid the formulation and solution of our SDP relaxations, we implement our techniques in QUINOPT, an open-source add-on to the MATLAB optimization toolbox YALMIP. We demonstrate our techniques by solving typical problems that arise in the context of stability analysis for dynamical systems governed by PDEs.

Key words. Integral inequalities, semidefinite programming, sum-of-squares optimization.

1. Introduction

Dynamical systems governed by partial differential equations (PDEs) are ubiquitous in physics and engineering. Being able to characterize their dynamics is essential to understand the underlying physical processes and, ultimately, control them to achieve a desired state. PDEs generally pose additional challenges compared to systems of ordinary differential equations (ODEs) 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).

A common approach to reduce the complexity of the analysis is to consider spatially-averaged quantities by means of volume integrals, which in some sense

*This work was partly supported by the Engineering and Physical Sciences Research Council (EPSRC) grant EP/J010537/1. Email addresses: gf910@ic.ac.uk (Giovanni Fantuzzi), a.wynn@imperial.ac.uk (Andrew Wynn), paul.goulart@eng.ox.ac.uk (Paul Goulart), antonis@eng.ox.ac.uk (Antonis Papachristodoulou).

†Department of Aeronautics, Imperial College London, London, UK.

‡Department of Engineering Science, University of Oxford, Oxford, UK.

allow one to account for the spatially-distributed dynamics using a single time-dependent variable. For example, the stability of an equilibrium solution of a PDE defined over a domain Ω can be established by constructing a positive integral Lyapunov functional $\mathcal{V}(t) = \mathcal{V}\{\boldsymbol{w}(t,\cdot)\} = \int_\Omega V[\boldsymbol{w}(t,\boldsymbol{x})]d\boldsymbol{x}$ whose time derivative is non-positive [25, 28, 29]. More recently, it has been shown that the input-to-state/output properties of some well-posed PDEs (e.g. their passivity, reachability, input-to-state stability) can be studied using dissipation inequalities for integral functionals of the state variable [2, 3]. Finally, bounds on time- and space-averaged properties for non-linear dissipative systems such as turbulent flows can be computed with optimization problems that involve integral inequality constraints [7, 8, 10–12, 18].

As highlighted by the previous examples, working with integral quantities often leads to the problem of determining whether a certain integral functional of the system state, say $\mathcal{F}\{\boldsymbol{w}\}$, is sign-definite for all admissible states. Alternatively, the problem is to determine values of the system parameters that make $\mathcal{F}\{\boldsymbol{w}\}$ sign-definite while optimizing a given cost function. Such problems can typically be written in the form

$$\begin{aligned} \min_{\boldsymbol{\gamma}} \quad & \boldsymbol{c}^T\boldsymbol{\gamma} \\ \text{subject to} \quad & \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \int_\Omega F_{\boldsymbol{\gamma}}(\boldsymbol{x}, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}) dx_1 \cdots dx_n \geq 0 \quad \forall \boldsymbol{w} \in H. \end{aligned} \tag{1}$$

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

When the dependence on $\boldsymbol{\gamma}$ is (at least) affine and strong duality holds, problem (1) could be solved by first computing the minimizer $\boldsymbol{w}^\star$ of $\mathcal{F}_{\boldsymbol{\gamma}}$ as a function of $\boldsymbol{\gamma}$ using the calculus of variations [9, 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 constraint. This strategy has been successfully applied to some problems in fluid dynamics (see e.g. [13, 31, 32]), but it generally 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 techniques [5]. More recently, it has been suggested that problems of type (1) can be recast as SDPs even when the integrand is a polynomial of $\boldsymbol{x}$ and $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$ [22, 28–30]. The idea is to relate the derivatives of the components of $\boldsymbol{w}$ using integration by parts and algebraic identities, then require 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.

In this paper, we develop a third approach to solve a class of problems of type (1), generalizing previous work by some of the authors [15, 16]. Specifically, we derive finite-dimensional inner and outer approximations of the feasible set of (1),

$$T := \{\boldsymbol{\gamma} \in \mathbb{R}^s : \ \forall \boldsymbol{w} \in H, \ \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq 0\}\,, \tag{2}$$

when $\mathcal{F}_\gamma$ is a homogeneous quadratic functional 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 $F_\gamma(\boldsymbol{x}, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w})$ is a homogeneous quadratic polynomial with respect to $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$. Using Legendre series expansions, we show that upper and lower bounds for the optimal value of (1) can be computed using semidefinite programming. Although our approach resembles the methods of [22, 28–30] in so far as it utilizes semidefinite programming, our strategy to formulate the SDP and the SDP itself differ fundamentally.

The rest of the paper is organized as follows. Sect. 2 describes the class of optimization problems studied in this work. We bound the optimal cost from below via convergent outer SDP relaxations in Sect. 3. Upper bounds are derived in Sect. 4 by constructing inner approximations of the feasible set T in terms of LMIs and SOS constraints. We remove some simplifying assumptions and further extend our results in Sect. 5. In Sect. 6 we present QUINOPT, an add-on to the MATLAB optimization toolbox YALMIP [20, 21] to aid the formulation of our SDP relaxations, and use it to solve some problems arising from the analysis of PDEs. Finally, Sect. 7 offers concluding remarks and perspectives for future developments.

Notation. Vectors and matrices are denoted by boldface characters. In particular, $\mathbf{0}$ denotes the zero vector/matrix; its size will be indicated if not clear from the context. Given a vector $\boldsymbol{v} \in \mathbb{R}^n$, $\|\boldsymbol{v}\|$ and $\|\boldsymbol{v}\|_1$ denote the usual Euclidean and ℓ^1 norms,

$$\|\boldsymbol{v}\| = \left(\sum_{i=1}^{n} |v_i|^2\right)^{1/2}, \qquad \|\boldsymbol{v}\|_1 = \sum_{i=1}^{n} |v_i|.$$

Given a matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times m}$, the Frobenius norm is defined as

$$\|\boldsymbol{Q}\|_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{m} |Q_{ij}|^2\right)^{1/2}.$$

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, i.e.,

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

The set of non-negative integers is denoted by $\mathbb{N}$, and $\mathbb{N}^q$ is the set of multi-indices of the form $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_q]$. 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 $i \in \{1, \dots, q\}$, we define $\boldsymbol{\beta} - \boldsymbol{\alpha} = [\beta_1 - \alpha_1, \dots, \beta_q - \alpha_q] \in \mathbb{N}^q$ and we list all multi-index derivatives of order between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in the vector

$$\mathcal{D}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w} := \left[\partial^{\alpha_1}u_1, \dots, \partial^{\beta_1}u_1, \partial^{\alpha_2}u_2, \dots, \partial^{\beta_2}u_2, \dots, \partial^{\beta_q}u_q\right]^T \in \mathbb{R}^{q+|\boldsymbol{\beta}-\boldsymbol{\alpha}|}. \tag{3}$$

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

$$\mathcal{B}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w} := \begin{bmatrix} \mathcal{D}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w}(a) \\ \mathcal{D}^{[\boldsymbol{\alpha},\boldsymbol{\beta}]}\boldsymbol{w}(b) \end{bmatrix} \in \mathbb{R}^{2(q+|\boldsymbol{\beta}-\boldsymbol{\alpha}|)}. \tag{4}$$

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

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.

2. Optimization with affine homogeneous quadratic integral inequalities

Let $\boldsymbol{\gamma} \in \mathbb{R}^s$ be a vector of optimization variables, and consider 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

$$1 \le k_i \le m-1, \qquad i \in \{1, \ldots, q\}, \tag{5a}$$
$$k_i \le l_i \le m, \qquad i \in \{1, \ldots, q\}. \tag{5b}$$

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

$$\boldsymbol{F}(x;\boldsymbol{\gamma}) := \boldsymbol{F}_0(x) + \sum_{i=1}^{s} \gamma_i \boldsymbol{F}_i(x). \tag{6}$$

In other words, $\boldsymbol{F}(x;\boldsymbol{\gamma})$ is a symmetric matrix of polynomials of x of degree at most d_F, the coefficients of which depend affinely on $\boldsymbol{\gamma}$.

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

$$\begin{aligned} \min_{\boldsymbol{\gamma}} \quad & \boldsymbol{c}^T\boldsymbol{\gamma} \\ \text{subject to} \quad & \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \int_{-1}^{1} \left(\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\right)^T \boldsymbol{F}(x;\boldsymbol{\gamma})\, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\, dx \ge 0, \quad \forall \boldsymbol{w} \in H, \end{aligned} \tag{7}$$

where $\boldsymbol{c} \in \mathbb{R}^s$ is the cost vector, $\boldsymbol{F}(x;\boldsymbol{\gamma})$ is as in (6), and

$$H := \left\{\boldsymbol{w} \in C^m\left([-1,1],\mathbb{R}^q\right) : \ \boldsymbol{A}\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} = \mathbf{0}\right\} \tag{8}$$

is the space of m-times continuously differentiable functions satisfying p homogeneous BCs defined by the matrix $\boldsymbol{A} \in \mathbb{R}^{p\times 2(q+|\boldsymbol{l}|)}$. Note that there is no loss of generality in fixing the integration domain for the functional $\mathcal{F}_{\boldsymbol{\gamma}}$ to $[-1,1]$ because a compact interval $[a,b]$ can be mapped to it with a change of integration variable. It can be checked that an affine homogeneous quadratic integral inequality represents a convex constraint on $\boldsymbol{\gamma}$, which makes (7) a convex optimization problem.

Remark 1. We allow the space H to be defined by derivatives of higher order than those appearing in $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ for generality (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 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.

Simplifying assumption. To facilitate the exposition, we will hereafter consider two-dimensional functions $\boldsymbol{w} = [u, v]^T \in C^m([-1, 1], \mathbb{R}^2)$. Moreover, we will restrict the attention to the uniform multi-indices $\boldsymbol{k} = [k, k]$ and $\boldsymbol{l} = [l, l]$, where the integers k and l satisfy (5a) and (5b). As will be discussed in Sect. 5, however, all our results hold for the general case.

Motivating Example. Consider a two-dimensional 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 non-dimensional magnitude 0.5γ, as shown in Figure 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 [16, 18, 27]. This steady flow is stable when the driving stress is small. The critical value $\gamma_{\rm cr}$ 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

$$\begin{aligned} &\min \quad -\gamma \\ \text{s.t.} \quad \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} := \int_{-1}^{1} &\left\{ \frac{16}{\xi^2}[(\partial_y^2 u)^2 + (\partial_y^2 v)^2] + 8[(\partial_y u)^2 + (\partial_y v)^2] \right. \\ &\left. + \xi^2(u^2 + v^2) + \frac{2\gamma}{\xi}(v\partial_y u - u\partial_y v) \right\} dy \geq 0, \end{aligned} \tag{9}$$

where the integral inequality constraint should hold for all functions $[u, v]^T$ satisfying the homogeneous BCs

$$u(-1) = u(1) = \partial_y u(-1) = \partial_y^2 u(1) = v(-1) = v(1) = \partial_y v(-1) = \partial_y^2 v(1) = 0. \tag{10}$$

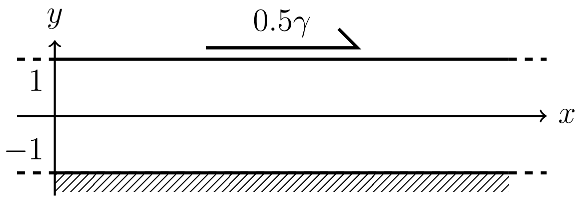

Figure 1: Sketch of the flow setup in our motivating example. The two-dimensional fluid layer extends to infinity along the x direction, is bounded at $y = -1$ by a solid boundary and is driven at the surface ($y = 1$) by a shear stress of non-dimensional magnitude 0.5γ.

See [18, 27] for a detailed discussion. The constraint in (9) can be rewritten in matrix form as in (7) with $\boldsymbol{k} = \boldsymbol{l} = [2, 2]$ and

$$\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w} = \begin{bmatrix} u \\ \partial_y u \\ \partial_y^2 u \\ v \\ \partial_y v \\ \partial_y^2 v \end{bmatrix}, \qquad \boldsymbol{F}(x;\boldsymbol{\gamma}) = \begin{bmatrix} \xi^2 & 0 & 0 & 0 & -\frac{\gamma}{\xi} & 0 \\ 0 & 8 & 0 & \frac{\gamma}{\xi} & 0 & 0 \\ 0 & 0 & \frac{16}{\xi^2} & 0 & 0 & 0 \\ 0 & \frac{\gamma}{\xi} & 0 & \xi^2 & 0 & 0 \\ -\frac{\gamma}{\xi} & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{16}{\xi^2} \end{bmatrix}. \tag{11}$$

Note that the matrix $\boldsymbol{F}$ above can be written in the form (6) 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 exact representation 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 non-trivial to solve. We will compute upper and lower bounds for the optimal γ in Sect. 6.1.

3. Outer SDP relaxations

Our first approach to solve (7) is to derive a sequence of *outer* approximations for its feasible set T in (2). In other words, we look for a family of sets $\{T_N^{\text{out}}\}_{N\geq 0}$ such that $T \subset T_N^{\text{out}}$. Optimizing the cost function over T_N^{out} then gives a lower bound for the optimal value of (7).

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

$$\boldsymbol{w} = [u, v]^T \in S_N := H \cap (\mathcal{P}_N \times \mathcal{P}_N) \subset H, \tag{12}$$

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 non-empty for any degree bound N because H contains the zero polynomial, and it contains nonzero elements if N is large enough to guarantee a sufficient number of degrees of freedom to satisfy the BCs prescribed on H in (8). Finally, $S_N \subset S_{N+1}$ since $\mathcal{P}_N \subset \mathcal{P}_{N+1}$.

Now, let $\check{\boldsymbol{u}}_N = [\hat{u}_0, \ldots, \hat{u}_N]^T$ and $\check{\boldsymbol{v}}_N = [\hat{v}_0, \ldots, \hat{v}_N]^T$ be the vectors of coefficients representing the polynomials u and v in any chosen basis for $\mathcal{P}_N$, and define $\boldsymbol{\psi}_N := \left[\check{\boldsymbol{u}}_N^T, \check{\boldsymbol{v}}_N^T\right]^T$. Since $\mathcal{F}_{\boldsymbol{\gamma}}$ in (7) 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

$$\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} = {\boldsymbol{\psi}_N}^T \boldsymbol{Q}_N(\boldsymbol{\gamma}) \boldsymbol{\psi}_N, \tag{13}$$

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

$$\boldsymbol{w} \in S_N \Leftrightarrow \boldsymbol{A}_N \boldsymbol{\psi}_N = \boldsymbol{0}. \tag{14}$$

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

$$\begin{aligned} T_N^{\text{out}} &:= \{\boldsymbol{\gamma} \in \mathbb{R}^s : \ \forall \boldsymbol{w} \in S_N,\ \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\} \geq 0\} \\ &= \left\{\boldsymbol{\gamma} \in \mathbb{R}^s : \ {\boldsymbol{\Pi}_N}^T \boldsymbol{Q}_N(\boldsymbol{\gamma}) \boldsymbol{\Pi}_N \succeq 0\right\}, \end{aligned} \tag{15}$$

and since $S_N \subset S_{N+1} \subset H$, the feasible set T of (7), defined as in (2), satisfies

$$T \subset T_{N+1}^{\text{out}} \subset T_N^{\text{out}}, \quad N \in \mathbb{N}. \tag{16}$$

This suggests that a sequence of lower bounds on the optimal value of (7) can be found by solving a series of truncated optimization problems.

Theorem 1. *Let $\boldsymbol{\gamma}^*$ be the optimal solution of (7) and, for each integer N, let $\boldsymbol{\gamma}_N^*$ be the optimal solution of the SDP*

$$\begin{aligned} \min_{\boldsymbol{\gamma}} \quad & \boldsymbol{c}^T \boldsymbol{\gamma} \\ \textit{subject to} \quad & \boldsymbol{\Pi}_N^T \boldsymbol{Q}_N(\boldsymbol{\gamma}) \, \boldsymbol{\Pi}_N \succeq 0. \end{aligned} \tag{17}$$

Then, $\{\boldsymbol{c}^T \boldsymbol{\gamma}_N^\star\}_{N \geq 0}$ is a non-decreasing sequence of lower bounds for $\boldsymbol{c}^T \boldsymbol{\gamma}^\star$. Moreover, $\lim_{N \to \infty} \|\boldsymbol{\gamma}_N^\star - \boldsymbol{\gamma}^\star\| = 0$.

Proof. See Appendix B.1. □

It is important to note that although Theorem 1 implies that the sequence of lower bounds converges, it provides no control on the gap $|\boldsymbol{c}^T(\boldsymbol{\gamma}_N^\star - \boldsymbol{\gamma}^\star)|$ between the cost of the full and truncated optimization problems *as a function of* N. In other words, an arbitrarily large N might be required for a given level of approximation accuracy. Consequently, in the rest of this work we focus on proving checkable conditions upon which upper bounds can be placed on $\boldsymbol{c}^T \boldsymbol{\gamma}^*$.

Remark 2. In the statement of Theorem 1 we have tacitly assumed that (7) and the SDP (17) have well defined solutions. No difficulties arise if (7) is either unbounded below or infeasible: in the former case, (17) is also unbounded below for any choice of N; in the latter case, the optimal value of (17) tends to infinity as N is increased. In particular, the infeasibility of the SDP (17) for a certain N provides a *certificate of infeasibility* for (7). However, we emphasize that the feasibility (resp. unboundedness) of (17) for any finite N does *not* prove that (7) is feasible (resp. unbounded).

4. Inner SDP relaxations

Upper bounds on the optimal value of (7) 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 (8). This strategy is complementary to the approach followed in Sect. 3, 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 non-negativity 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 (7), and the corresponding cost $\boldsymbol{c}^T \boldsymbol{\gamma}$ is an upper bound for the optimal value of (7).

4.1. Legendre series expansions

The key to derive a lower bound $\mathcal{F}_\gamma\{\boldsymbol{w}\} \geq \mathcal{G}_\gamma\{\boldsymbol{w}\}$ which is valid for all $\boldsymbol{w} = [u, v]^T \in H$ (recall our simplifying restriction to the two-dimensional case) is to expand u and v in terms of Legendre polynomials. That is, we write expansions such as

$$\partial^\alpha u = \sum_{n=0}^{\infty} \hat{u}_n^\alpha \mathcal{L}_n(x), \tag{18}$$

where $\mathcal{L}_n(x)$ is the Legendre polynomial of degree n and $\hat{u}_n^\alpha$ is the n-th 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., $\int_{-1}^{1} \mathcal{L}_m \mathcal{L}_n \, dx = 0$ if $m \neq n$ [19]. This will enable us to enforce the non-negativity of the functional $\mathcal{F}_\gamma$ in (7) 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 are only orthogonal with respect to a weighting function, and so they do not suit our purposes. A short introduction to Legendre polynomials, Legendre series and their properties is given in Appendix A; see [1, 19, 33] for a comprehensive treatment of the subject.

To avoid working with infinite series and to facilitate our analysis, we decompose (18) into a finite sum and a remainder function. More precisely, given an integer i we define the remainder function

$$U_i^\alpha(x) = \sum_{n=i+1}^{\infty} \hat{u}_n^\alpha \mathcal{L}_n(x). \tag{19}$$

Next, we choose an integer N such that

$$N \geq d_F + k - 1, \tag{20}$$

where d_F is the degree of the polynomial matrix $\boldsymbol{F}$ defined in (6). For each $\alpha \in \{1, \ldots, k\}$ we decompose the Legendre expansion of $\partial^\alpha u$ as

$$\partial^\alpha u = \sum_{n=0}^{N+\alpha} \hat{u}_n^\alpha \, \mathcal{L}_n(x) + U_{N+\alpha}^\alpha(x). \tag{21}$$

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

$$\hat{\boldsymbol{u}}_{[r,s]}^\alpha = \begin{bmatrix} \hat{u}_r^\alpha, & \ldots, & \hat{u}_s^\alpha \end{bmatrix}^T \in \mathbb{R}^{s-r+1}. \tag{22}$$

For technical reasons that will be pointed out in Sect. 4.2, it will also be convenient to introduce an "extended" decomposition for the highest-order derivative, $\partial^k u$. Specifically, we let

$$M := N + 2k + d_F \tag{23}$$

and consider

$$\partial^k u = \sum_{n=0}^{M} \hat{u}_n^k \, \mathcal{L}_n(x) + U_M^k(x). \tag{24}$$

The following result, proven in Appendix B.2, illustrates how the Legendre coefficients of u, ∂u, $\ldots$, $\partial^k u$ can be related.

Lemma 1. *Let $u \in C^m([-1,1])$ and its derivatives up to order $k \leq m-1$ be expanded as in (21), and let M be as in (23). For any $\alpha \in \{1, \dots, k\}$ and any two integers r, s with $0 \leq r \leq s \leq M+\alpha-k$, there exist matrices $\boldsymbol{B}^{\alpha}_{[r,s]}$ and $\boldsymbol{D}^{\alpha}_{[r,s]}$ such that*

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

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

This lemma simply states that given the Legendre coefficients $\hat{u}^k_0, \dots, \hat{u}^k_M$ 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

$$\check{\boldsymbol{u}}_n = \begin{bmatrix} \mathcal{D}^{k-1}u(-1) \\ \hat{u}^k_0 \\ \vdots \\ \hat{u}^k_n \end{bmatrix} \in \mathbb{R}^{k+n+1}. \tag{25}$$

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 (7) is only required to hold for functions that satisfy prescribed BCs. The following result is proven in Appendix B.3.

Lemma 2. *Let $u \in C^m([-1,1])$ and its derivatives up to order $k \leq m-1$ be expanded as in (21), and let $\mathcal{B}^{k-1}u \in \mathbb{R}^{2k}$ be defined according to (4). Moreover, let M be as in (23), and let $\check{\boldsymbol{u}}_M \in \mathbb{R}^{k+M+1}$ be defined according to (25). There exists a matrix $\boldsymbol{G}_M \in \mathbb{R}^{2k\times(k+M+1)}$ such that*

$$\mathcal{B}^{k-1}u = \boldsymbol{G}_M \check{\boldsymbol{u}}_M.$$

4.2. Legendre expansions of $\mathcal{F}_\gamma\{\boldsymbol{w}\}$

Recalling the definition of $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}$, we see from (7) that $\mathcal{F}_\gamma\{\boldsymbol{w}\}$ is a sum of elementary terms of the form

$$\int_{-1}^{1} f\, \partial^\alpha u\, \partial^\beta v\, dx, \tag{26}$$

where $\alpha, \beta \in \{0, \dots, k\}$. 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 (26), we substitute $\partial^\alpha u$ and $\partial^\beta v$ with their decomposed Legendre expansions according to the following strategy:

- If $\alpha \neq k$ *or* $\beta \neq k$, use (21).
- If $\alpha = \beta = k$, use the "extended" decomposition (24).

In either case, we can rewrite (26) as

$$\int_{-1}^{1} f\, \partial^\alpha u\, \partial^\beta v\, dx = \mathcal{P}_{uv}^{\alpha\beta} + \mathcal{Q}_{uv}^{\alpha\beta} + \mathcal{R}_{uv}^{\alpha\beta}, \tag{27}$$

where

$$\mathcal{P}_{uv}^{\alpha\beta} = \sum_{m=0}^{N_\alpha} \sum_{n=0}^{N_\beta} \hat{u}_m^\alpha \hat{v}_n^\beta \int_{-1}^{1} f\, \mathcal{L}_m\, \mathcal{L}_n\, dx, \tag{28a}$$

$$\mathcal{Q}_{uv}^{\alpha\beta} = \sum_{n=0}^{N_\alpha} \hat{u}_n^\alpha \int_{-1}^{1} f\, \mathcal{L}_n\, V_{N_\beta}^\beta\, dx + \sum_{n=0}^{N_\beta} \hat{v}_n^\beta \int_{-1}^{1} f\, \mathcal{L}_n\, U_{N_\alpha}^\alpha\, dx, \tag{28b}$$

$$\mathcal{R}_{uv}^{\alpha\beta} = \int_{-1}^{1} f\, U_{N_\alpha}^\alpha\, V_{N_\beta}^\beta\, dx. \tag{28c}$$

Here and in the following it should be understood that $N_\alpha = N+\alpha$ and $N_\beta = N+\beta$ when (21) is used to expand $\partial^\alpha u$ and $\partial^\beta v$, while $N_\alpha = N_\beta = M = N + 2k + d_F$ when (24) is used. We consider more Legendre coefficients for the derivatives of order k because Lemma 1 will be used to rewrite $\mathcal{P}_{uv}^{\alpha\beta}$, $\mathcal{Q}_{uv}^{\alpha\beta}$ and $\mathcal{R}_{uv}^{\alpha\beta}$ for all values of α and β in terms of a common set of variables.

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

$$\boldsymbol{\psi}_M := \begin{bmatrix} \check{\boldsymbol{u}}_M \\ \check{\boldsymbol{v}}_M \end{bmatrix} \in \mathbb{R}^{2(k+M+1)}, \tag{29}$$

where $\check{\boldsymbol{u}}_M$ and $\check{\boldsymbol{v}}_M$ are as in (25), we arrive at the following result.

Lemma 3. *Let $\mathcal{P}_{uv}^{\alpha\beta}$ be as in (28a) and $\boldsymbol{\psi}_M$ be defined according to (29). 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 will show in Appendix B.4 that the extended decomposition (24) 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 (28b) 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

$$\boldsymbol{L}_n\, \boldsymbol{\psi}_M = \begin{bmatrix} \hat{\boldsymbol{u}}_{[n,M]}^k \\ \hat{\boldsymbol{v}}_{[n,M]}^k \end{bmatrix}, \quad n \in \{0, \ldots, M\}, \tag{30}$$

and $\boldsymbol{L}_n\, \boldsymbol{\psi}_M = \boldsymbol{0}$ if $n > M$. The existence of $\boldsymbol{L}_n$ follows from (29), (25), and (22). 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 ij-th element is defined as

$$\left(\boldsymbol{\Phi}_{[a,b]}^{[c,d]}\right)_{ij} = \int_{-1}^{1} f\, \mathcal{L}_{m_i}\, \mathcal{L}_{n_j}\, dx, \tag{31}$$

where m_i and n_j are the i-th and j-th elements of the sequences $\{a, \dots, b\}$ and $\{c, \dots, d\}$. Note that, strictly speaking, $\boldsymbol{\Phi}_{[c,d]}^{[a,b]}$ depends on f, and its entries are affine on $\boldsymbol{\gamma}$. We do not indicate such dependencies explicitly to avoid complicating our notation further. The following result is proven in Appendix B.4.

Lemma 4. *Let $\mathcal{Q}_{uv}^{\alpha\beta}$ be as in (28b) and let d_F be the degree of $f(x;\boldsymbol{\gamma})$.*

(i) 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

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

(ii) If $\alpha = \beta = k$, define $\overline{M} := M + 1 - d_F$. Moreover, let

$$\boldsymbol{\Delta} := \operatorname{Diag}\left(\frac{2}{2(M+1)+1}, \dots, \frac{2}{2(M+d_F)+1}\right) \in \mathbb{S}^{d_F}, \tag{32a}$$

$$\boldsymbol{Y}(\boldsymbol{\gamma}) := \frac{1}{2}\begin{bmatrix} \boldsymbol{0} & \boldsymbol{\Phi}_{[M+1-d_F,M]}^{[M+1,M+d_f]} \\ \boldsymbol{\Phi}_{[M+1-d_F,M]}^{[M+1,M+d_f]} & \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^{2d_F \times 2d_F}. \tag{32b}$$

Finally, let $\boldsymbol{R}_{uv}^{kk} \in \mathbb{S}^{2d_F}$ and a diagonal matrix $\boldsymbol{\Sigma}_{uv}^{kk} \in \mathbb{S}^2$ satisfy the LMI

$$\boldsymbol{\Omega}(\boldsymbol{R}_{uv}^{kk}, \boldsymbol{\Sigma}_{uv}^{kk}, \boldsymbol{\gamma}) := \begin{bmatrix} \boldsymbol{R}_{uv}^{kk} & \boldsymbol{Y}(\boldsymbol{\gamma}) \\ \boldsymbol{Y}(\boldsymbol{\gamma})^T & \boldsymbol{\Sigma}_{uv}^{kk} \otimes \boldsymbol{\Delta} \end{bmatrix} \succeq 0, \tag{33}$$

where $\otimes$ is the usual Kronecker product. Then

$$\mathcal{Q}_{uv}^{kk} \geq -{\boldsymbol{\psi}_M}^T \left({\boldsymbol{L}_{\overline{M}}}^T \boldsymbol{R}_{uv}^{kk} \boldsymbol{L}_{\overline{M}}\right) \boldsymbol{\psi}_M - \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \boldsymbol{\Sigma}_{uv}^{kk} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} dx. \tag{34}$$

Remark 3. The estimate (34) is a generalized version of Young's inequality. To mitigate the fact that such estimates are typically conservative, we consider the matrices $\boldsymbol{R}_{uv}^{kk}$ and $\boldsymbol{\Sigma}_{uv}^{kk}$ as auxiliary optimization variables, which should be determined subject to (33) in order to make (34) as sharp as possible.

Lemmas 3 and 4 show that the terms $\mathcal{P}_{uv}^{\alpha\beta}$ and $\mathcal{Q}_{uv}^{\alpha\beta}$ can be bounded in terms of $\boldsymbol{\psi}_M$, U_M^k and V_M^k for any $\alpha, \beta \in \{0, \dots, k\}$. If $\alpha = \beta = k$, (28c) is also written in terms of U_M^k and V_M^k. The following result, proven in Appendix B.5, shows that $\mathcal{R}_{uv}^{\alpha\beta}$ can be estimated in terms of the same quantities also when $\alpha \neq k$ or $\beta \neq k$.

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

$$\mathcal{R}_{uv}^{\alpha\beta} \geq -\|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 {\boldsymbol{\psi}_M}^T \boldsymbol{R}_{uv}^{\alpha\beta} \boldsymbol{\psi}_M - \|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \boldsymbol{\Sigma}_{uv}^{\alpha\beta} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} dx. \tag{35}$$

4.3. A lower bound for $\mathcal{F}_{\gamma}\{\boldsymbol{w}\}$

Let us now combine Lemmas 3–5 to find a lower bound for the integral functional $\mathcal{F}_{\gamma}\{\boldsymbol{w}\}$ in (7). 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 (7) corresponding to the entries $\partial^k u$ and $\partial^k v$ of $\mathcal{D}^k\boldsymbol{w}$. The contribution of the terms with $\alpha = \beta = k$ to $\mathcal{F}_{\gamma}\{\boldsymbol{w}\}$ is

$$\int_{-1}^{1} \begin{bmatrix}\partial^k u\\ \partial^k v\end{bmatrix}^T \boldsymbol{S}(x;\boldsymbol{\gamma}) \begin{bmatrix}\partial^k u\\ \partial^k v\end{bmatrix} \, dx. \tag{36}$$

It follows from Lemma 3 and part (ii) of Lemma 4 that

$$\begin{aligned}\int_{-1}^{1} \begin{bmatrix}\partial^k u\\ \partial^k v\end{bmatrix}^T \boldsymbol{S}(x;\boldsymbol{\gamma}) \begin{bmatrix}\partial^k u\\ \partial^k v\end{bmatrix} \, dx \geq {\boldsymbol{\psi}_M}^T \left(\boldsymbol{P}_{uu}^{kk} + 2\boldsymbol{P}_{uv}^{kk} + \boldsymbol{P}_{vv}^{kk}\right) \boldsymbol{\psi}_M \\ - {\boldsymbol{\psi}_M}^T {\boldsymbol{L}_{M+1-d_F}}^T \left(\boldsymbol{R}_{uu}^{kk} + 2\boldsymbol{R}_{uv}^{kk} + \boldsymbol{R}_{vv}^{kk}\right) \boldsymbol{L}_{M+1-d_F}\, \boldsymbol{\psi}_M \\ + \int_{-1}^{1} \begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix}^T \left[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_{uu}^{kk} - 2\boldsymbol{\Sigma}_{uv}^{kk} - \boldsymbol{\Sigma}_{vv}^{kk}\right] \begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix} dx,\end{aligned} \tag{37}$$

where the auxiliary variables $\boldsymbol{R}_{uu}^{kk}$, $\boldsymbol{\Sigma}_{uu}^{kk}$, $\boldsymbol{R}_{uv}^{kk}$, $\boldsymbol{\Sigma}_{uv}^{kk}$, $\boldsymbol{R}_{vv}^{kk}$, and $\boldsymbol{\Sigma}_{vv}^{kk}$ must satisfy three LMIs defined as in (33). For notational convenience, we let

$$\mathcal{Y} = \left\{\boldsymbol{R}_{uu}^{kk}, \boldsymbol{\Sigma}_{uu}^{kk}, \boldsymbol{R}_{uv}^{kk}, \boldsymbol{\Sigma}_{uv}^{kk}, \boldsymbol{R}_{vv}^{kk}, \boldsymbol{\Sigma}_{vv}^{kk}\right\} \tag{38}$$

be the list of all auxiliary variables, and we combine the three LMIs they must satisfy into the equivalent block-diagonal LMI

$$\overline{\boldsymbol{\Omega}}(\boldsymbol{\gamma}, \mathcal{Y}) := \begin{bmatrix}\boldsymbol{\Omega}(\boldsymbol{R}_{uu}^{kk}, \boldsymbol{\Sigma}_{uu}^{kk}, \boldsymbol{\gamma}) & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{\Omega}(\boldsymbol{R}_{uv}^{kk}, \boldsymbol{\Sigma}_{uv}^{kk}, \boldsymbol{\gamma}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\Omega}(\boldsymbol{R}_{vv}^{kk}, \boldsymbol{\Sigma}_{vv}^{kk}, \boldsymbol{\gamma})\end{bmatrix} \succeq 0. \tag{39}$$

All terms contributing to $\mathcal{F}_{\gamma}\{\boldsymbol{w}\}$ with $\alpha \neq k$ or $\beta \neq k$ can instead be lower bounded using Lemmas 3–5 to obtain expressions of the form

$$\begin{aligned}\int_{-1}^{1} f\, \partial^{\alpha} u\, \partial^{\beta} v\, dx \geq {\boldsymbol{\psi}_M}^T \left(\boldsymbol{P}_{uv}^{\alpha\beta} + \boldsymbol{Q}_{uv}^{\alpha\beta} - \|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 \boldsymbol{R}_{uv}^{\alpha\beta}\right) \boldsymbol{\psi}_M \\ - \|\hat{\boldsymbol{f}}(\boldsymbol{\gamma})\|_1 \int_{-1}^{1} \begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix}^T \boldsymbol{\Sigma}_{uv}^{\alpha\beta} \begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix} dx.\end{aligned} \tag{40}$$

From (37) and (40) 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

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

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

4.4. Projection onto the boundary conditions

The lower bound (41) holds for any continuously differentiable function $\boldsymbol{w}$, irrespectively of whether it satisfies the BCs prescribed on H. Recalling (8), these are given by the set of p homogeneous equations

$$\boldsymbol{A}\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} = \boldsymbol{0}. \tag{42}$$

To enforce as many BCs as possible in (41) and to sharpen the lower bound over the space H, we need to rewrite (42) in terms of our Legendre expansions.

We begin by introducing a permutation matrix $\boldsymbol{P}$ such that

$$\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} = \boldsymbol{P}\begin{bmatrix}\mathcal{B}^{\boldsymbol{k}-\boldsymbol{1}}\boldsymbol{w}\\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}\end{bmatrix}. \tag{43}$$

so (42) becomes

$$\boldsymbol{A}\boldsymbol{P}\begin{bmatrix}\mathcal{B}^{\boldsymbol{k}-\boldsymbol{1}}\boldsymbol{w}\\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}\end{bmatrix} = \boldsymbol{0}. \tag{44}$$

A straightforward corollary of Lemma 2 and (29) is that there exists a matrix $\boldsymbol{J}$ such that $\mathcal{B}^{\boldsymbol{k}-\boldsymbol{1}}\boldsymbol{w} = \boldsymbol{J}\boldsymbol{\psi}_M$. Then, (44) can be rewritten as

$$\boldsymbol{K}\begin{bmatrix}\boldsymbol{\psi}_M\\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}\end{bmatrix} = \boldsymbol{0}, \quad \boldsymbol{K} := \boldsymbol{A}\boldsymbol{P}\begin{bmatrix}\boldsymbol{J} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}. \tag{45}$$

Any solution of (45) can be expressed as

$$\begin{bmatrix}\boldsymbol{\psi}_M\\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Lambda}_1\\ \boldsymbol{\Lambda}_2\end{bmatrix}\boldsymbol{\theta} =: \boldsymbol{\Lambda}\boldsymbol{\theta} \tag{46}$$

for some $\boldsymbol{\theta} \in \mathbb{R}^{\dim[\mathcal{N}(\boldsymbol{K})]}$, where $\boldsymbol{\Lambda}$ is a matrix satisfying $\mathcal{R}(\boldsymbol{\Lambda}) = \mathcal{N}(\boldsymbol{K})$. In general, the submatrices $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ may have linearly dependent columns, so it is possible to write $\boldsymbol{\psi}_M$ and $\mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w}$ with the more "economical" representation

$$\boldsymbol{\psi}_M = \boldsymbol{\Pi}_M\boldsymbol{\zeta}, \qquad \mathcal{R}(\boldsymbol{\Pi}_M) = \mathcal{R}(\boldsymbol{\Lambda}_1), \quad \boldsymbol{\zeta} \in \mathbb{R}^{\dim[\mathcal{R}(\boldsymbol{\Lambda}_1)]}, \tag{47a}$$

$$\mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} = \boldsymbol{\Gamma}\boldsymbol{\eta}, \qquad \mathcal{R}(\boldsymbol{\Gamma}) = \mathcal{R}(\boldsymbol{\Lambda}_2), \quad \boldsymbol{\eta} \in \mathbb{R}^{\dim[\mathcal{R}(\boldsymbol{\Lambda}_2)]}. \tag{47b}$$

Next, (47b) can be turned into a set of BCs for the remainder functions U_M^k and V_M^k, parametrized by $\boldsymbol{\zeta}$ and $\boldsymbol{\eta}$. Differentiating (24) yields

$$\partial^{k+i}u(x) = \sum_{n=0}^{M-i}\hat{c}_n^i(\boldsymbol{\psi}_M)\mathcal{L}_n(x) + \partial^i U_M^k(x), \quad i \in \{0, \ldots, l-k\}, \tag{48}$$

where the coefficients $\hat{c}_n^i$ are known linear functions of $\boldsymbol{\psi}_M$. Similar expressions can be written for $\partial^{k+i}v(x)$. After evaluating such expressions at $x = \pm 1$, we can therefore find a vector-valued linear function $\boldsymbol{h}(\boldsymbol{\psi}_M)$ such that

$$\mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} = \boldsymbol{h}(\boldsymbol{\psi}_M) + \mathcal{B}^{\boldsymbol{l}-\boldsymbol{k}}\begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix}. \tag{49}$$

Using (47a) and (47b), we conclude that U_M^k and V_M^k should satisfy the BCs

$$\mathcal{B}^{\boldsymbol{l}-\boldsymbol{k}}\begin{bmatrix}U_M^k\\ V_M^k\end{bmatrix} = \boldsymbol{\Gamma}\boldsymbol{\eta} - \boldsymbol{h}\left(\boldsymbol{\Pi}_M\boldsymbol{\zeta}\right). \tag{50}$$

Substituting (47a) into (41), we finally conclude that when (39) holds, $\mathcal{F}_\gamma\{\boldsymbol{w}\}$ can be lower bounded over the space H in (8) as

$$\mathcal{F}_\gamma\{\boldsymbol{w}\} \geq \boldsymbol{\zeta}^T \, \boldsymbol{\Pi}_M^T \boldsymbol{Q}_M \boldsymbol{\Pi}_M \, \boldsymbol{\zeta} + \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \left[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M \right] \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} dx, \tag{51}$$

where the remainder functions U_M^k and V_M^k are orthogonal to all Legendre polynomials of degree less than or equal to M and satisfy the BCs (50).

4.5. Formulating an inner SDP relaxation

The integral inequality in (7) is satisfied if the right-hand side of (51) is non-negative for all $\boldsymbol{\zeta}, \boldsymbol{\eta}$ and all admissible remainder functions U_M^k and V_M^k. Recalling that (51) is valid only if (39) holds, we have the following result.

Proposition 1. *Let $M = M(N)$ be as in (23) for any integer N, and let $\mathcal{Y}$ be as in (38). 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*

$$\overline{\boldsymbol{\Omega}}(\boldsymbol{\gamma};\mathcal{Y}) \succeq 0, \tag{52a}$$

$$\boldsymbol{\Pi}_M^T \, \boldsymbol{Q}_M(\boldsymbol{\gamma},\mathcal{Y}) \, \boldsymbol{\Pi}_M \succeq 0, \tag{52b}$$

$$\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y}) \geq 0, \quad \forall x \in [-1,1], \tag{52c}$$

is an inner approximation to the feasible set T of (7), i.e., $T_N^{\text{in}} \subset T$.

Conditions (52b) and (52c) are only sufficient, not necessary, for the non-negativity of the right-hand side of (51): they do not take into account the boundary and orthogonality conditions on the remainder functions. However, they are useful because they can be turned into tractable constraints.

For example, (52b) 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 (6) 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 $-t \leq \hat{f}_n(\boldsymbol{\gamma}) \leq t$ [6].

Moreover, (52c) is an LMI if the matrix $\boldsymbol{S}(x;\boldsymbol{\gamma})$ is independent of x, which is true in many interesting and non-trivial cases, such as our motivating example in Sect. 2. Otherwise, (52c) is equivalent to the polynomial inequality

$$\boldsymbol{z}^T[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y})]\boldsymbol{z} \geq 0, \quad \forall (x,\boldsymbol{z}) \in [-1,1] \times \mathbb{R}^2. \tag{53}$$

Although checking a polynomial inequality is generally NP-hard (see [23, Sect. 2.1] and references therein), we can turn (53) into an LMI plus linear equality constraints by a SOS relaxation [23]. Using the so-called $\mathcal{S}$-procedure [26], we introduce a tunable symmetric polynomial matrix $\boldsymbol{T}(x) \in \mathbb{S}^2$ and require that the multivariate polynomials

$$p_1(x,\boldsymbol{z}) := \boldsymbol{z}^T[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y}) - (1-x^2)\boldsymbol{T}(x)]\boldsymbol{z}, \tag{54a}$$

$$p_2(x,\boldsymbol{z}) := \boldsymbol{z}^T\boldsymbol{T}(x)\boldsymbol{z} \tag{54b}$$

admit a SOS decomposition; it is not difficult to see that this implies (53).

This suggests that an upper bound for the optimal value of (7), as well as a feasible point that achieves it, can be found by solving an SDP.

Theorem 2. *Let $M = M(N)$ be defined as in (23) for any integer N, let $\mathcal{Y}$ be as in (38), and let $\boldsymbol{T}(x) \in \mathbb{S}^2$ be a tunable polynomial matrix. The optimal value of the SDP*

$$\begin{aligned}
\min_{\boldsymbol{\gamma},\mathcal{Y},\boldsymbol{T}(x)} \quad & \boldsymbol{c}^T\boldsymbol{\gamma}, \\
\textit{subject to} \quad & \overline{\boldsymbol{\Omega}}(\boldsymbol{\gamma};\mathcal{Y}) \succeq 0, \\
& \boldsymbol{\Pi}_M^T \, \boldsymbol{Q}_M(\boldsymbol{\gamma},\mathcal{Y}) \, \boldsymbol{\Pi}_M \succeq 0, \\
& \boldsymbol{z}^T \left[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M(\boldsymbol{\gamma},\mathcal{Y}) - (1-x^2)\boldsymbol{T}(x) \right] \boldsymbol{z} \quad \textit{is SOS}, \\
& \boldsymbol{z}^T\boldsymbol{T}(x)\boldsymbol{z} \quad \textit{is SOS},
\end{aligned} \tag{55}$$

is an upper bound for the optimal value of (7). Moreover, the corresponding minimizer $\boldsymbol{\gamma}_N^\star$ is a feasible point for (7).

Remark 4. In contrast to our results for the outer SDP relaxations of Sect. 3, we cannot prove that the optimal solution of (55) converges to that of the original problem as N is increased, nor that the optimal value of (55) is non-increasing. In fact, without further assumptions on the functional $\mathcal{F}_{\boldsymbol{\gamma}}$ in (7), it is possible that (55) is always infeasible even if (7) is feasible. To see this, recall that the matrix $\boldsymbol{\Sigma}_M$ is positive definite, so (52c) 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

$$\int_{-1}^{1} \left[x^2(\partial u)^2 + (\partial v)^2 - \gamma uv \right] \, dx \geq 0, \tag{56}$$

which can be written in the form (7) with $\boldsymbol{k} = (1,1)$; for simplicity, we do not prescribe any boundary conditions, which amounts to setting $\boldsymbol{A} = \boldsymbol{0}$ in (8). This inequality is clearly feasible for $\gamma = 0$. Yet, (55) 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 method of [22, 28–30] we could establish that (56) is feasible for approximately $|\gamma| \leq 1.2$. With the exception of such pathological cases, however, our inner SDP relaxations are observed to work well in practice; we demonstrate this in Sect. 6. 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.

5. Extensions

5.1. Inequalities with explicit dependence on boundary values

In the applications we have in mind, the integral inequality constraint in (7) is derived from a weak formulation of a PDE, after integrating some terms by parts. Sometimes, the BCs are such that the boundary terms from such integrations by parts do not vanish; we will see an example in Sect. 6.2.

This motivates us to extend our results to quadratic homogeneous functionals that depend explicitly on the boundary values $\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}$, i.e.

$$\mathcal{F}_{\gamma}\{\boldsymbol{w}\} := \int_{-1}^{1} \Big[\left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{bnd}}(x;\boldsymbol{\gamma})\, \mathcal{B}^{\boldsymbol{l}}\boldsymbol{w} + \left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{mix}}(x;\boldsymbol{\gamma})\, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w} \\ + \left(\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{int}}(x;\boldsymbol{\gamma})\, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w} \Big]\, dx, \quad (57)$$

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 (6). Note that (57) reduces to the functional in (7) if one sets $\boldsymbol{F}_{\text{mix}} = \boldsymbol{0}$, $\boldsymbol{F}_{\text{bnd}} = \boldsymbol{0}$ 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 (43). Upon integrating the known matrix $\boldsymbol{P}^T\boldsymbol{F}_{\text{bnd}}(x;\boldsymbol{\gamma})\boldsymbol{P}$, it follows from (29) and Lemma 2 that there exists a symmetric matrix $\boldsymbol{Q}_M^{\text{bnd}}(\boldsymbol{\gamma})$ such that

$$\int_{-1}^{1} \left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{bnd}}(x;\boldsymbol{\gamma})\, \mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\, dx = \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \boldsymbol{Q}_M^{\text{bnd}}(\boldsymbol{\gamma}) \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}. \quad (58)$$

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 (20) that we have decomposed the Legendre expansion of $\partial^\alpha u$ with the truncation parameter $N \geq d_F + k - 1$, we can write

$$\begin{aligned} \int_{-1}^{1} g_i(x;\boldsymbol{\gamma})\partial^\alpha u\, dx &= \sum_{m=0}^{d_F}\sum_{n=0}^{\infty} \hat{g}_{i,m}(\boldsymbol{\gamma})\hat{u}_n^\alpha \int_{-1}^{1} \mathcal{L}_m\, \mathcal{L}_n\, dx \\ &= \left[2\hat{g}_{i,0}(\boldsymbol{\gamma}), \quad \frac{2\hat{g}_{i,1}(\boldsymbol{\gamma})}{3}, \quad \ldots, \quad \frac{2\hat{g}_{i,d_F}(\boldsymbol{\gamma})}{2d_F+1} \right] \hat{\boldsymbol{u}}^\alpha_{[0,d_F]} \\ &= \left[2\hat{g}_{i,0}(\boldsymbol{\gamma}), \quad \frac{2\hat{g}_{i,1}(\boldsymbol{\gamma})}{3}, \quad \ldots, \quad \frac{2\hat{g}_{i,d_F}(\boldsymbol{\gamma})}{2d_F+1} \right] \left[\boldsymbol{B}^\alpha_{[0,d_F]} \quad \boldsymbol{D}^\alpha_{[0,d_F]}\right] \check{\boldsymbol{u}}_M. \end{aligned} \quad (59)$$

The last equality follows from Lemma 1 after defining $\check{\boldsymbol{u}}_M$ according to (25). With the help of (29) and Lemma 2 it is then possible to find a matrix $\boldsymbol{Q}_M^{\text{mix}}(\boldsymbol{\gamma})$ such that

$$\begin{aligned} \int_{-1}^{1} \left(\mathcal{B}^{\boldsymbol{l}}\boldsymbol{w}\right)^T \boldsymbol{F}_{\text{mix}}(x;\boldsymbol{\gamma})\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\, dx &= \begin{bmatrix} \mathcal{B}^{\boldsymbol{k}-\boldsymbol{1}}\boldsymbol{w} \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \int_{-1}^{1} \boldsymbol{P}^T\boldsymbol{F}_{\text{mix}}(x;\boldsymbol{\gamma})\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}\, dx \\ &= \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \boldsymbol{Q}_M^{\text{mix}}(\boldsymbol{\gamma})\, \boldsymbol{\psi}_M. \end{aligned} \quad (60)$$

Note that (58) and (60) are exact formulae, and no approximation is made. Combining these results with (41), we conclude that there is a symmetric matrix $\boldsymbol{Q}_M^{\text{tot}} = \boldsymbol{Q}_M^{\text{tot}}(\boldsymbol{\gamma}, \mathcal{Y})$ such that

$$\mathcal{F}_{\gamma}\{\boldsymbol{w}\} \geq \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix}^T \boldsymbol{Q}_M^{\text{tot}} \begin{bmatrix} \boldsymbol{\psi}_M \\ \mathcal{B}^{[\boldsymbol{k},\boldsymbol{l}]}\boldsymbol{w} \end{bmatrix} + \int_{-1}^{1} \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix}^T \left[\boldsymbol{S}(x;\boldsymbol{\gamma}) - \boldsymbol{\Sigma}_M\right] \begin{bmatrix} U_M^k \\ V_M^k \end{bmatrix} dx. \quad (61)$$

Finally, we can use (46) to enforce the BCs, and we conclude that Proposition 1 and Theorem 2 hold when we replace (52b) with

$$\boldsymbol{\Lambda}^T\boldsymbol{Q}_M^{\text{tot}}(\boldsymbol{\gamma})\boldsymbol{\Lambda} \succeq 0. \quad (62)$$

5.2. Higher-dimensional function spaces & generic multi-index derivatives

Theorems 1 and 2 were derived under 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 Sect. 5.1, 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 (5a) and (5b).

In particular, all our proofs extend verbatim by simply identifying the functions u, v used throughout Sect. 3 and 4 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 non-uniform multi-indices $\boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}^q$ requires only minor modifications; the details are left to the interested reader.

6. 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 our SDP relaxations, we have developed QUINOPT (QUadratic INtegral OPTimization), an open-source add-on for the MATLAB optimization toolbox YALMIP [20, 21]. QUINOPT uses the Legendre polynomial basis for the outer SDP relaxations of Sect. 3, because the orthogonality of the Legendre polynomials promotes sparsity of the SDP data. QUINOPT, the scripts used to produce the results in the following sections, and additional examples can be downloaded from

https://github.com/giofantuzzi/QUINOPT/releases.

Our experiments were run on a PC with a 3.40GHz Intel® Core™ i7-4770 CPU and 16Gb of RAM, using MOSEK [4] to solve our SDP relaxations.

6.1. Motivating example: stability of a stress-driven shear flow

Consider our motivating example of Sect. 2, 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 [16]; here, we replicate those results using our general-purpose toolbox QUINOPT. Since we minimize the negative of γ in (9), the SDPs (55) and (17) give, respectively, lower and upper bounds for the stress $\gamma_{\rm cr}$ at which the flow is no longer provably stable.

Figure 2 shows the upper and lower bounds for $\gamma_{\rm cr}$ as a function of the wave number ξ, a parameter in (9), computed for three 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 (10), and (17) reduces to an unconstrained minimization problem yielding an infinite upper bound. More detailed numerical results, along with CPU times, are reported in Table 1 for two values of the wave number parameter, $\xi = 3$ and $\xi = 9$.

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 1). This means that we can bound $\gamma_{\rm cr}$ accurately and extremely efficiently using our techniques. Moreover, it suggests that the inner SDP relaxations converge to the full problem (9), despite our inability to provide a proof of this fact in general (cf. Remark 4).

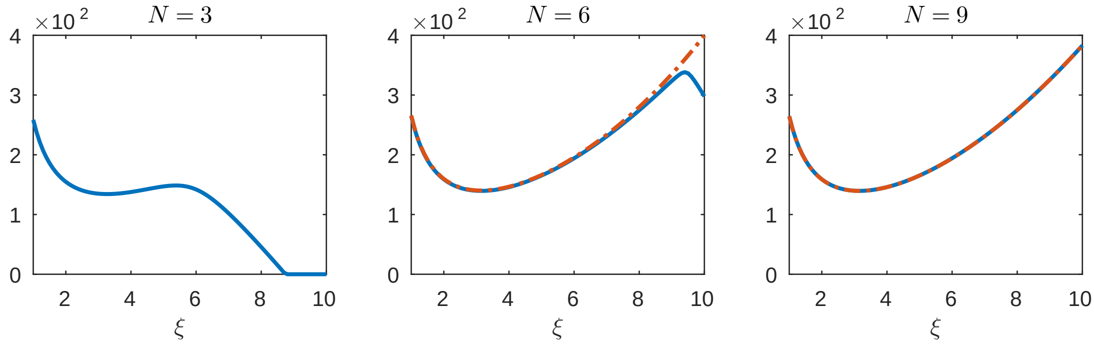

Figure 2: (Color online) Upper (dot-dashed line) and lower (solid line) bounds on the optimal value of (9) as a function of ξ for different values of N. The upper bound for $N = 3$ is infinite and so it is not plotted. The bounds are indistinguishable for $N = 9$.

Table 1: Selected upper bounds (UB) and lower bounds (LB) for $\gamma_{\rm cr}$ and CPU times ($t_{\rm LB}$ and $t_{\rm UB}$) in seconds for different values of the Legendre truncation parameter N.

N	$\xi = 3$				$\xi = 9$			
	LB	$t_{\rm LB}$	UB	$t_{\rm UB}$	LB	$t_{\rm LB}$	UB	$t_{\rm UB}$
3	134.8594	0.09	+INF	0.03	0.0000	0.08	+INF	0.03
6	139.7656	0.10	140.4087	0.04	323.5764	0.08	335.1022	0.04
9	139.7700	0.08	139.7701	0.06	325.6449	0.09	325.6764	0.05
12	139.7700	0.10	139.7700	0.05	325.6453	0.10	325.6455	0.05

6.2. Stability of a system of coupled PDEs

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

$$\partial_t \boldsymbol{w} = \gamma \partial_x^2 \boldsymbol{w} + \boldsymbol{A}\boldsymbol{w}, \qquad \boldsymbol{A} = \begin{bmatrix} 1 & 1.5 \\ 5 & 0.2 \end{bmatrix}, \tag{63}$$

over the domain $[0,1]$, subject to the BCs $u(0) = u(1) = v(0) = v(1) = 0$. This system was studied in [28, Sect. V.D] with the different 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 $\gamma < \gamma_{\rm cr} = 0.341216$. Since the system is linear, then it is stable with respect to finite-amplitude perturbations for all $\gamma \geq \gamma_{\rm cr}$.

Following [28], we try to establish the stability of the system with respect to arbitrary perturbations by considering Lyapunov functionals of the form

$$\mathcal{V}(t) = \frac{1}{2}\int_0^1 \boldsymbol{w}^T \boldsymbol{P}(x)\boldsymbol{w}\, dx, \tag{64}$$

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 $-\frac{d\mathcal{V}}{dt} \geq 0$. 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 (63) to compute $\frac{d\mathcal{V}}{dt}$, we find that the critical value of γ at which (64) stops being a valid Lyapunov function for a given degree d_P is given by

$$\begin{aligned} \min_{\gamma, \boldsymbol{P}(x)} \quad & \gamma \\ \text{subject to} \quad & \int_0^1 \boldsymbol{w}^T \boldsymbol{P}(x) \boldsymbol{w} \, dx \geq 0, \\ & \int_0^1 \boldsymbol{w}^T \boldsymbol{P}(x)(-\gamma \partial_x^2 \boldsymbol{w} - \boldsymbol{A}\boldsymbol{w}) \, dx \geq 0. \end{aligned} \tag{65}$$

The optimization variables are γ and the coefficients of the entries of $\boldsymbol{P}(x)$, so the problem is not jointly convex in γ and $\boldsymbol{P}$. To avoid non-convexity, we either fix $\boldsymbol{P}$ *a priori* and minimize γ, or fix γ and try to find a suitable $\boldsymbol{P}$ by solving a feasibility problem; the minimum γ at which a feasible $\boldsymbol{P}$ stops existing can be determined with a bisection method.

Before deriving our SDP relaxations, we need to rescale the domain of integration for the constraints in (65) to $[-1, 1]$. Moreover, in light of Remark 4, 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 after rescaling and integration by parts the second integral inequality in (65) depends explicitly on the unspecified boundary values $\partial_x u(\pm 1)$ and $\partial_x v(\pm 1)$, making the extensions discussed in Sect. 5.1 necessary.

Table 2 shows the upper and lower bounds for the optimal solution of (65) as a function of the degree d_P, obtained with the SDPs (55) and (17) respectively. We also considered the particular choice $\boldsymbol{P}(x) = \boldsymbol{I}$, which corresponds to the classical approach of taking the energy of the system as the candidate Lyapunov function. In all cases, we fixed the Legendre series truncation parameter to $N = 10$ and the degree of the matrix $\boldsymbol{T}(x)$ in (55) to 6, which gives well converged results. The average CPU time taken by QUINOPT to set up and solve each feasibility problem in our bisection procedure, or to minimize γ when $\boldsymbol{P}(x) = \boldsymbol{I}$, is also shown.

Our results show that stability can be established up to the known critical value $\gamma_{\text{cr}} = 0.341216$ for all choices of d_P, with the exception of the classical energy Lyapunov function. This drastically improves the conservative results obtained in [28] for the same problem, reported in Table 2 for comparison (the original results are for a parameter $R = \gamma^{-1}$ and have been adapted). Our results demonstrate that our SDP relaxations accurately approximate (65); this is particularly significant for

Table 2: Upper bounds (UB) and lower bounds (LB) for the optimal solution of (65) for different choices of d_P. The upper bounds from [28] are also reported (note that the original results are for a parameter $R = \gamma^{-1}$). The average CPU time (t_{LB} and t_{UB}) required to set up and solve each feasibility problem in our bisection procedure is also shown.

d_P	UB from [28]	UB	t_{UB}	LB	t_{LB}
0 ($\boldsymbol{P}(x) = \boldsymbol{I}$)	5	0.392571	0.2670	0.392571	0.0858
0	3.333333	0.341216	0.1497	0.341216	0.1211
2	0.588235	0.341216	1.3257	0.341216	0.9874
4	0.434783	0.341216	1.5709	0.341216	1.0731
6	0.416667	0.341216	1.8222	0.341216	1.1756
8	0.408163	0.341216	2.4731	0.341216	1.3501

the inner SDP approximations, which rely on typically conservative estimates and for which we cannot prove convergence.

6.3. Feasible set approximation

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

$$\int_{-1}^{1} \left[(\partial u)^2 + (\partial v)^2 + \gamma_1 x^2 \partial u\, \partial v + 2\gamma_2 uv\right] dx \geq 0, \tag{66}$$

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.

Outer and inner approximation sets T_N^{out} and T_N^{in} can be found using the SDPs (17) and (55), respectively. In particular, we compute the boundaries of T_N^{out} and T_N^{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 the SOS optimization problem (55), we fix the degree of the tunable polynomial matrix $\boldsymbol{T}(x)$ to 6; our results do not improve when this is increased.

Figure 3 shows the difference between T_N^{out} and T_N^{in} for eight different choices of the Legendre series truncation parameter N; note that $N = 2$ is the minimum value that satisfies (20). The difference between T_N^{out} and T_N^{in} reduces visibly as N is increased from 2 to 16, but it seems to stabilize for $N \geq 20$. One possible cause for this is a very slow convergence of the outer approximation sets T_N^{out} to the feasible set of (66) when N is large. The most likely reason, however, is that the inner approximation sets do not converge to the true feasible set; in fact, our estimates in Lemmas 4 and 5 and the SOS relaxation of the polynomial matrix inequality (52c) can be expected to introduce some conservativeness.

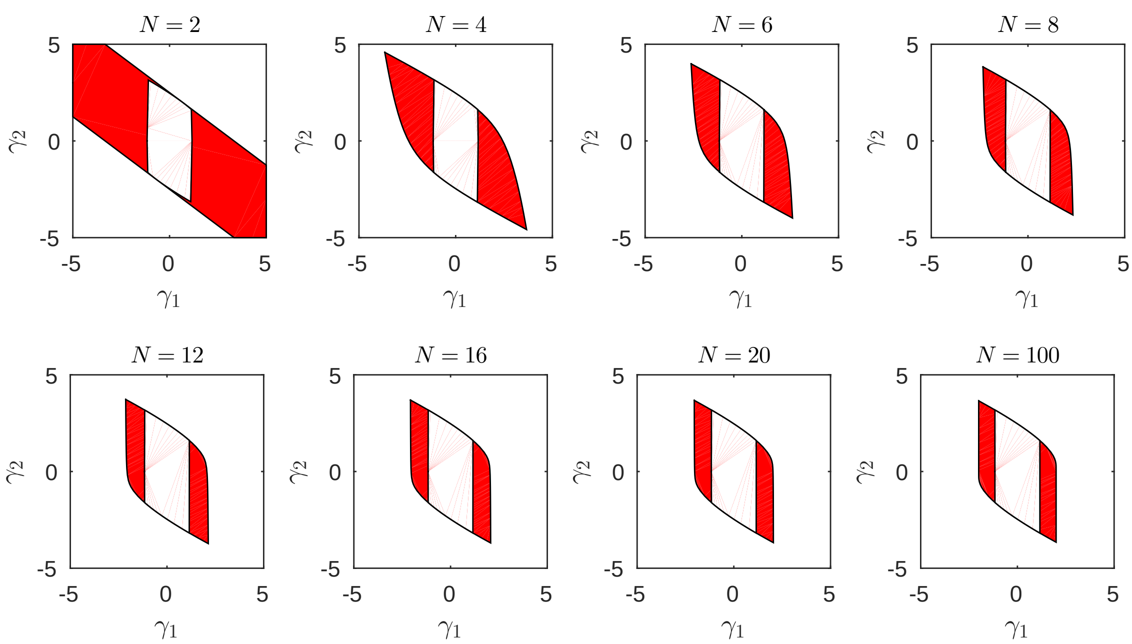

Figure 3: (Color online) Difference between inner and outer approximation of the feasible set of (66) for different choices of Legendre series truncation parameter N.

Interestingly, however, some parts of the boundaries of T_N^{out} and T_N^{in} almost coincide, even for N as low as 4. In particular, the figures suggest that the inner approximation sets T_N^{in} are only over-constrained in the γ_1 direction. This is because γ_2 only appears in the term $\int_{-1}^{1} 2\gamma_2 uv\,dx$, 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 $\int_{-1}^{1} \gamma_1 x^2 \partial u\,\partial v\,dx$, to which the estimates in part (ii) of Lemma 4 must be applied. Despite our efforts to tune the auxiliary matrices in (34), 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.

7. Conclusion

In this work, 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 non-decreasing sequence that converges to the exact optimal cost value. On the other hand, we have seen that 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 some non-trivial problems that arise when studying the stability of dynamical systems governed by 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. Based on the results presented in Sect. 6.3, we expect that more stringent assumption on the properties of the integral inequality would have to be introduced.

Looking at the applications we have in mind, i.e., the analysis of systems governed by PDEs, the present work should be extended to integral inequalities over two and higher dimensional domains. For compact box domains, this can be achieved simply by introducing Legendre expansions in each coordinate direction and adapting the ideas presented in this work. For domains with a more general shape, including non-compact domains, other basis functions could be used. This may pose additional challenges in the derivation of inner approximations, because they require estimates that rely on specific properties of the basis functions.

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 Sect. 5.1. 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 [3] are to be successfully applied to complex nonlinear systems of PDEs of interest in physics and engineering.

A. Introduction to Legendre polynomials and Legendre series

The following material is a summary of important notions regarding Legendre polynomials and Legendre series expansions taken from Refs. [1, 19, 33].

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

$$\mathcal{L}_n(x) = \frac{1}{n!\,2^n}\frac{d^n}{dx^n}(x^2-1)^n. \tag{67}$$

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

$$(n+1)\,\mathcal{L}_{n+1}(x) = (2n+1)\,x\,\mathcal{L}_n(x) - n\mathcal{L}_{n-1}(x), \tag{68}$$

with $\mathcal{L}_0(x) = 1$ and $\mathcal{L}_1(x) = x$. Equation (68) 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

$$(2n+1)\mathcal{L}_n(x) = \frac{d}{dx}\left[\mathcal{L}_{n+1}(x) - \mathcal{L}_{n-1}(x)\right], \quad n \geq 1, \tag{69}$$

see e.g. [1, 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 $L^2(-1,1)$ [33], and satisfy the orthogonality condition

$$\int_{-1}^{1} \mathcal{L}_n\,\mathcal{L}_m\,dx = \frac{2\delta_{mn}}{2n+1}, \tag{70}$$

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)

$$u(x) = \sum_{n=0}^{\infty} \hat{u}_n \mathcal{L}_n(x), \quad \hat{u}_n = \int_{-1}^{1} u\,\mathcal{L}_n\,dx, \tag{71}$$

where the $\hat{u}_n$'s are known as Legendre coefficients. From (70) it follows that

$$\|u\|_2^2 = \int_{-1}^{1} |u|^2\,dx = \sum_{n=0}^{\infty} \frac{2|\hat{u}_n|^2}{2n+1}. \tag{72}$$

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

$$|u(y) - u(x)| = |\partial u(z)|\,|x-y| \le \|\partial u\|_\infty\,|x-y| \le C\,|x-y|.$$

Here, C is a generic positive constant whose existence is guaranteed by the continuity of ∂u in $[-1,1]$. The uniform convergence property is established in light of [19, Theorem XI and subsequent comments].

B. Proofs

B.1. Proof of Theorem 1

The sequence of optimal values $\{\boldsymbol{c}^T\boldsymbol{\gamma}_{\boldsymbol{N}}^\star\}_{N\ge 0}$ is non-decreasing since $T_{N+1}^{\mathrm{out}} \subset T_N^{\mathrm{out}}$. The convergence of $\boldsymbol{\gamma}_N^*$ to $\boldsymbol{\gamma}^*$ as $N \to \infty$ follows from the fact that the set T_N^{out}, i.e., the feasible set of the SDP (17), converges to the feasible set T of (7) as $N \to \infty$.

By convergence to the feasible set, we mean that if $\boldsymbol{\gamma}$ is infeasible for (7), it is also infeasible for the SDP relaxation (17) when N is sufficiently large. To prove this, define

$$t(\boldsymbol{\gamma}) := \inf_{\boldsymbol{w}\in H} \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}, \qquad t_N(\boldsymbol{\gamma}) := \inf_{\boldsymbol{w}\in S_N} \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}, \tag{73}$$

so T and T_N^{out} are described, respectively, by the inequalities $t(\boldsymbol{\gamma}) \ge 0$ and $t_N(\boldsymbol{\gamma}) \ge 0$. We need not assume that these infima are achieved.

Lemma B.1. *For any $\varepsilon > 0$, for each $\boldsymbol{\gamma} \in \mathbb{R}^s$ there exists $N = N(\varepsilon, \boldsymbol{\gamma}) \in \mathbb{N}$ such that*

$$0 \le t_N(\boldsymbol{\gamma}) - t(\boldsymbol{\gamma}) \le \varepsilon.$$

Proof. We prove the equivalent result that $t_N(\boldsymbol{\gamma}) - t(\boldsymbol{\gamma}) \downarrow 0$ as $N \to \infty$. Let $\{\boldsymbol{w}_n\}_{n\ge 0}$ with $\boldsymbol{w}_n \in H$ be a minimizing sequence for $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$ over H, i.e., such that

$$\lim_{n\to\infty} \mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}_n\} = t(\boldsymbol{\gamma}). \tag{74}$$

The continuity of the integrand of $\mathcal{F}_{\boldsymbol{\gamma}}\{\boldsymbol{w}\}$, i.e.,

$$F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)\right) = \left(\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)\right)^T \boldsymbol{F}(x;\boldsymbol{\gamma})\,\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x), \tag{75}$$

with respect to all entries of the vector $\mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)$ at a fixed $x \in [-1,1]$ implies that there exists $\delta_n(x) > 0$ such that for all $\boldsymbol{w} \in H$

$$\max_{0\le\alpha\le k} |\partial^\alpha \boldsymbol{w}(x) - \partial^\alpha \boldsymbol{w}_n(x)|_\infty \le \delta_n(x) \Rightarrow \\ \left|F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)\right) - F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)\right)\right| \le \frac{1}{2n}. \tag{76}$$

Here $\partial^\alpha \boldsymbol{w}(x)$ means that the differential operator ∂^α applies to each component of $\boldsymbol{w}$, and we use $|\cdot|_\infty$ to distinguish the ℓ^∞ norm of a vector from the L^∞ norm of a function.

Letting δ_n be the minimum of $\delta_n(x)$ over $[-1,1]$, and noting that

$$|\mathcal{F}_\gamma\{\boldsymbol{w}\} - \mathcal{F}_\gamma\{\boldsymbol{w}_n\}| \leq 2 \left\| F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)\right) - F\left(x, \mathcal{D}^{\boldsymbol{k}}\boldsymbol{w}(x)\right) \right\|_\infty, \tag{77}$$

it follows from (76) that

$$\max_{0\leq\alpha\leq k} \|\partial^\alpha \boldsymbol{w} - \partial^\alpha \boldsymbol{w}_n\|_\infty \leq \delta_n \Rightarrow |\mathcal{F}_\gamma\{\boldsymbol{w}\} - \mathcal{F}_\gamma\{\boldsymbol{w}_n\}| \leq \frac{1}{n}. \tag{78}$$

Since the Weierstrass approximation theorem can be extended to linear subspaces of continuously differentiable functions with prescribed boundary conditions (this follows e.g from [24, Proposition 2]), there exists a polynomial $\boldsymbol{P}_n \in H$ of degree d_n that satisfies (78). Without loss of generality, we may also assume that $d_n < d_{n+1}$.

Now, for all $N \in \{d_n, \ldots, d_{n+1}-1\}$ we have that $\boldsymbol{P}_n \in S_N$, so

$$t(\boldsymbol{\gamma}) \leq t_N(\boldsymbol{\gamma}) \leq \mathcal{F}_\gamma\{\boldsymbol{P}_n\} \leq |\mathcal{F}_\gamma\{\boldsymbol{P}_n\} - \mathcal{F}_\gamma\{\boldsymbol{w}_n\}| + \mathcal{F}_\gamma\{\boldsymbol{w}_n\} \leq \frac{1}{n} + \mathcal{F}_\gamma\{\boldsymbol{w}_n\}. \tag{79}$$

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 $N \to \infty$. □

Now, suppose $\boldsymbol{\gamma} \in \mathbb{R}^s$ is not feasible for (7), i.e. $\boldsymbol{\gamma} \notin T$. Then,

$$t(\boldsymbol{\gamma}) \leq -2\varepsilon < 0 \tag{80}$$

for some $\varepsilon > 0$. By Lemma B.1, there exists an integer N such that

$$t_N(\boldsymbol{\gamma}) \leq \varepsilon + t(\boldsymbol{\gamma}) \leq \varepsilon - 2\varepsilon = -\varepsilon < 0, \tag{81}$$

which implies that $\boldsymbol{\gamma} \notin T_N^{\mathrm{out}}$. We conclude that the family of sets $\{T_N^{\mathrm{out}}\}_{N\geq 0}$ converges to T as $N \to 0$, and hence the optimal solution of (17) converges to that of (7).

B.2. 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 $\partial^\alpha u$, $\alpha \in \{0, ..., k\}$ converge uniformly, cf. Appendix A. Consequently, we can use the fundamental theorem of calculus for each $\alpha \leq k-1$ to write

$$(\partial^\alpha u)(x) = \partial^\alpha u(-1) + \int_{-1}^x \partial^{\alpha+1} u(t)\,dt = \partial^\alpha u(-1) + \sum_{n\geq 0} \hat{u}_n^{\alpha+1} \int_{-1}^x \mathcal{L}_n(t)\,dt. \tag{82}$$

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 (69). We can then rewrite (82) as

$$\partial^\alpha u = \quad \partial^\alpha u(-1) + [\mathcal{L}_1 + \mathcal{L}_0]\,\hat{u}_0^{\alpha+1} + \sum_{n\geq 1} [\mathcal{L}_{n+1} - \mathcal{L}_{n-1}]\,\frac{\hat{u}_n^{\alpha+1}}{2n+1}. \tag{83}$$

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

$$\hat{u}_0^\alpha = \partial^\alpha u(-1) + \hat{u}_0^{\alpha+1} - \frac{1}{3}\hat{u}_1^{\alpha+1}, \tag{84a}$$

$$\hat{u}_n^\alpha = \frac{\hat{u}_{n-1}^{\alpha+1}}{2n-1} - \frac{\hat{u}_{n+1}^{\alpha+1}}{2n+3}, \qquad n \geq 1. \tag{84b}$$

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

$$\hat{\boldsymbol{u}}^\alpha_{[r,s]} = \boldsymbol{E}^\alpha \mathcal{D}^{k-1}u(-1) + \boldsymbol{C}^\alpha \hat{\boldsymbol{u}}^{\alpha+1}_{[r-1,s+1]}. \tag{85}$$

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, it follows from (84) that $\boldsymbol{E}^\alpha = \mathbf{0}$ if $r \geq 1$.

Expressions similar to (85) can be built for all vectors $\hat{\boldsymbol{u}}^{\alpha+i}_{[r-i,s+i]}$, with $i \in \{0, \ldots, k-\alpha-1\}$. After some algebra, it is therefore possible to write

$$\hat{\boldsymbol{u}}^\alpha_{[r,s]} = \boldsymbol{B}^\alpha_{[r,s]} \mathcal{D}^{k-1}u(-1) + \left(\prod_{i=0}^{k-\alpha-1} \boldsymbol{C}^{\alpha+i} \right) \hat{\boldsymbol{u}}^k_{[r-k+\alpha,s+k-\alpha]}, \tag{86}$$

where

$$\boldsymbol{B}^\alpha_{[r,s]} := \boldsymbol{E}^\alpha + \boldsymbol{C}^\alpha \boldsymbol{E}^{\alpha+1} + \cdots + \left(\prod_{i=0}^{k-\alpha-2} \boldsymbol{C}^{\alpha+i} \right) \boldsymbol{E}^{k-1}. \tag{87}$$

Note that, in light of (84), 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 (86) 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 of Lemma 1 is concluded by defining

$$\boldsymbol{D}^\alpha_{[r,s]} := \left[\mathbf{0}_{(s-r+1)\times(r-k+\alpha)}, \quad \prod_{i=0}^{k-\alpha-1} \boldsymbol{C}^{\alpha+i}, \quad \mathbf{0}_{(s-r+1)\times(M-s-k+\alpha)} \right]. \tag{88}$$

B.3. 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.2, it may be shown that for any $\alpha \in \{0, \ldots, k-1\}$,

$$\partial^\alpha u(1) = \partial^\alpha u(-1) + \sum_{n=0}^{+\infty} \hat{u}^{\alpha+1}_n \int_{-1}^{1} \mathcal{L}_n(x) dx = \partial^\alpha u(-1) + 2\hat{u}^{\alpha+1}_0. \tag{89}$$

It then follows from Lemma 1 that $\partial^\alpha u(1)$ can 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 that the same is true for all entries of the vector $\mathcal{D}^{k-1}u(1)$, which proves the existence of $\boldsymbol{G}_M$.

B.4. Proof of Lemma 4

(i) Recall (19) and expand

$$\mathcal{Q}^{\alpha\beta}_{uv} = \sum_{m=0}^{N_\alpha} \sum_{n=N_\beta+1}^{+\infty} \hat{u}^\alpha_m \hat{v}^\beta_n \int_{-1}^{1} f\, \mathcal{L}_m\, \mathcal{L}_n\, dx + \sum_{m=N_\alpha+1}^{\infty} \sum_{n=0}^{N_\beta} \hat{u}^\alpha_m \hat{v}^\beta_n \int_{-1}^{1} f\, \mathcal{L}_m\, \mathcal{L}_n\, dx, \tag{90}$$

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 [14] that the integral $\int_{-1}^{1} f\, \mathcal{L}_n \mathcal{L}_m \, dx$ vanishes if $|m - n| > d_F$. Using the short-hand notation $\overline{n} = n + 1 - d_F$, we can write

$$\mathcal{Q}_{uv}^{\alpha\beta} = \begin{bmatrix} \hat{u}^{\alpha}_{\overline{N_\beta}} \\ \vdots \\ \hat{u}^{\alpha}_{N_\alpha} \end{bmatrix}^T \boldsymbol{\Phi}^{[N_\beta+1, N_\alpha+d_F]}_{[\overline{N_\beta}, N_\alpha]} \begin{bmatrix} \hat{v}^{\beta}_{N_\beta+1} \\ \vdots \\ \hat{v}^{\beta}_{N_\alpha+d_F} \end{bmatrix} + \begin{bmatrix} \hat{v}^{\beta}_{\overline{N_\alpha}} \\ \vdots \\ \hat{v}^{\beta}_{N_\beta} \end{bmatrix}^T \boldsymbol{\Phi}^{[N_\alpha+1, N_\beta+d_F]}_{[\overline{N_\alpha}, N_\beta]} \begin{bmatrix} \hat{u}^{\alpha}_{N_\alpha+1} \\ \vdots \\ \hat{u}^{\alpha}_{N_\beta+d_F} \end{bmatrix}. \tag{91}$$

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 (91) are well-defined. If the left (resp. right) inequality is not satisfied, then the first (resp. second) term in (91) 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 write

$$\mathcal{Q}_{uv}^{\alpha\beta} = \left(\hat{\boldsymbol{u}}^k_{[0,M]}\right)^T \boldsymbol{Q}(\boldsymbol{\gamma})\, \hat{\boldsymbol{v}}^k_{[0,M]} \tag{92}$$

for a suitably defined matrix $\boldsymbol{Q}(\boldsymbol{\gamma})$. The matrix $\boldsymbol{Q}^{\alpha\beta}_{uv}$ is found using (30) after taking the symmetric part of the right-hand side of (92).

(ii) Let

$$\boldsymbol{\nu} = \left[\hat{u}^k_{M+1}, \quad \ldots, \quad \hat{u}^k_{M+d_F}, \quad \hat{v}^k_{M+1}, \quad \ldots, \quad \hat{v}^k_{M+d_F}\right]^T. \tag{93}$$

After replacing N_α and N_β with M in (91), it may be verified using (30) that

$$\mathcal{Q}^{kk}_{uv} = 2\, {\boldsymbol{\psi}_M}^T\, {\boldsymbol{L}_{\overline{M}}}^T \boldsymbol{Y} \boldsymbol{\nu}. \tag{94}$$

Equation (33) then implies that

$$\begin{aligned} 0 &\leq \begin{bmatrix} \boldsymbol{L}_{\overline{M}} \boldsymbol{\psi}_M \\ \boldsymbol{\nu} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{R}^{kk}_{uv} & \boldsymbol{Y} \\ \boldsymbol{Y}^T & \boldsymbol{\Sigma}^{kk}_{uv} \otimes \boldsymbol{\Delta} \end{bmatrix} \begin{bmatrix} \boldsymbol{L}_{\overline{M}} \boldsymbol{\psi}_M \\ \boldsymbol{\nu} \end{bmatrix} \\ &= {\boldsymbol{\psi}_M}^T \left({\boldsymbol{L}_{\overline{M}}}^T \boldsymbol{R}^{kk}_{uv} \boldsymbol{L}_{\overline{M}}\right) \boldsymbol{\psi}_M + \boldsymbol{\nu}^T \left(\boldsymbol{\Sigma}^{kk}_{uv} \otimes \boldsymbol{\Delta}\right) \boldsymbol{\nu} + \mathcal{Q}^{kk}_{uv}. \end{aligned} \tag{95}$$

Now, $\boldsymbol{\Sigma}^{kk}_{uv}$ is a diagonal matrix by assumption. Recalling the definition of $\boldsymbol{\Delta}$ from (32a) and of $\boldsymbol{\nu}$ from (93), after some algebra we can rearrange (95) as

$$\begin{aligned} \mathcal{Q}^{kk}_{uv} \geq &-{\boldsymbol{\psi}_M}^T \left({\boldsymbol{L}_{\overline{M}}}^T \boldsymbol{R}^{kk}_{uv} \boldsymbol{L}_{\overline{M}}\right) \boldsymbol{\psi}_M \\ &- (\boldsymbol{\Sigma}^{kk}_{uv})_{11} \sum_{n=M+1}^{M+d_F} \frac{2|\hat{u}^k_n|^2}{2n+1} - (\boldsymbol{\Sigma}^{kk}_{uv})_{22} \sum_{n=M+1}^{M+d_F} \frac{2|\hat{v}^k_n|^2}{2n+1}. \end{aligned} \tag{96}$$

The proof is concluded by recognizing from (72) that the two sums in the last expression can be bounded, respectively, by $\left\|U^k_M\right\|_2^2$ and $\left\|V^K_M\right\|_2^2$.

B.5. 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 (19), (20) and (23), we can write

$$\frac{1}{2}\left\|U_{N_\alpha}^\alpha\right\|_2^2 = \sum_{n=N_\alpha+1}^{+\infty} \frac{(\hat{u}_n^\alpha)^2}{2n+1} = \sum_{n=N_\alpha+1}^{M-k+\alpha} \frac{(\hat{u}_n^\alpha)^2}{2n+1} + \sum_{n=M-k+\alpha+1}^{+\infty} \frac{(\hat{u}_n^\alpha)^2}{2n+1}$$
$$= \left(\hat{\boldsymbol{u}}_{[0,M]}^k\right)^T \boldsymbol{H}_\alpha\, \hat{\boldsymbol{u}}_{[0,M]}^k + \sum_{n=M-k+\alpha+1}^{+\infty} \frac{(\hat{u}_n^\alpha)^2}{2n+1}, \tag{97}$$

where the matrix $\boldsymbol{H}_\alpha$ can be obtained from Lemma 1. In particular, we note that (84b) is applied $k-a$ times to $(\hat{u}_n^\alpha)^2$ 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 (97) is exactly $\left\|U_M^k\right\|_2^2/2$ so

$$\frac{1}{2}\left\|U_{N_k}^k\right\|_2^2 = \left(\hat{\boldsymbol{u}}_{[0,M]}^k\right)^T \boldsymbol{H}_k\, \hat{\boldsymbol{u}}_{[0,M]}^k + \frac{1}{2}\left\|U_M^k\right\|_2^2. \tag{98}$$

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

$$\omega_\eta = \frac{4}{[2(M-k+\eta+1)-1][2(M-k+\eta+1)+3]}, \qquad \eta \in \{0, \ldots, k-1\}, \tag{99}$$

and use (84), the elementary inequality $(a-b)^2 \leq 2(a^2+b^2)$, and appropriate changes of indices to show that

$$\sum_{n=M-k+\alpha+1}^{+\infty} \frac{(\hat{u}_n^\alpha)^2}{2n+1} \leq \sum_{n=M-k+\alpha+1}^{+\infty} \frac{2}{2n+1}\left[\frac{|\hat{u}_{n-1}^{\alpha+1}|^2}{(2n-1)^2} + \frac{|\hat{u}_{n+1}^{\alpha+1}|^2}{(2n+3)^2}\right]$$
$$\leq \sum_{n=M-k+\alpha}^{M-k+\alpha+1} \frac{2|\hat{u}_n^{\alpha+1}|^2}{(2n+3)(2n+1)^2} + \sum_{n=M-k+\alpha+2}^{\infty} \frac{4|\hat{u}_n^{\alpha+1}|^2}{(2n-1)(2n+1)(2n+3)}$$
$$\leq \sum_{n=M-k+\alpha}^{M-k+\alpha+1} \frac{2|\hat{u}_n^{\alpha+1}|^2}{(2n+3)(2n+1)^2} + \omega_{\alpha+1} \sum_{n=M-k+\alpha+2}^{+\infty} \frac{|\hat{u}_n^{\alpha+1}|^2}{2n+1}. \tag{100}$$

Applying Lemma 1 to the first term on the right-hand side of (100) and substituting back into (97), we can construct a matrix $\boldsymbol{T}_\alpha$ such that

$$\frac{1}{2}\left\|U_{N_\alpha}^\alpha\right\|_2^2 \leq \left(\hat{\boldsymbol{u}}_{[0,M]}^k\right)^T \boldsymbol{T}_\alpha\, \hat{\boldsymbol{u}}_{[0,M]}^k + \omega_{\alpha+1} \sum_{n=M-k+\alpha+2}^{+\infty} \frac{|\hat{u}_n^{\alpha+1}|^2}{2n+1}. \tag{101}$$

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 (101). By recursion, we can eventually construct a matrix $\boldsymbol{Z}_\alpha$ and a constant λ_α such that

$$\frac{1}{2}\left\|U_{N_\alpha}^\alpha\right\|_2^2 \leq \left(\hat{\boldsymbol{u}}_{[0,M]}^k\right)^T \boldsymbol{Z}_\alpha\, \hat{\boldsymbol{u}}_{[0,M]}^k + \lambda_\alpha\left\|U_M^k\right\|_2^2. \tag{102}$$

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 (99). Moreover, the right-hand side of (102) has the same form as (98), so for the rest of this section we will not distinguish the cases $\alpha \leq k-1$ and $\alpha = k$.

The estimate (102) can be used in conjunction with Young's inequality and (30) to show that for any $\varepsilon > 0$ we can bound

$$|\mathcal{R}_{uv}^{\alpha\beta}| \leq \|f\|_\infty \boldsymbol{\psi}_M{}^T \left(\boldsymbol{L}_0{}^T \begin{bmatrix} \varepsilon \boldsymbol{Z}_\alpha & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\varepsilon}\boldsymbol{Z}_\beta \end{bmatrix} \boldsymbol{L}_0 \right) \boldsymbol{\psi}_M{}^T \\ + \|f\|_\infty + \left(\varepsilon\lambda_\alpha \|U_k\|_2^2 + \frac{\lambda_\beta}{\varepsilon} \|V_k\|_2^2 \right). \tag{103}$$

Setting $\varepsilon = (N+1)^{\beta-\alpha}$, so that $\varepsilon\lambda_\alpha \sim \varepsilon^{-1}\lambda_\beta \sim N^{\alpha+\beta-2k}$ and $\|\varepsilon \boldsymbol{Z}_\alpha\|_F \sim \|\varepsilon^{-1}\boldsymbol{Z}_\beta\|_F \sim N^{\alpha+\beta-2k-1}$, and letting

$$\boldsymbol{R}_{uv}^{\alpha\beta} := \boldsymbol{L}_0{}^T \begin{bmatrix} \varepsilon \boldsymbol{Z}_\alpha & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\varepsilon}\boldsymbol{Z}_\beta \end{bmatrix} \boldsymbol{L}_0, \qquad \boldsymbol{\Sigma}_{uv}^{\alpha\beta} := \begin{bmatrix} \varepsilon\lambda_\alpha & 0 \\ 0 & \varepsilon^{-1}\lambda_\beta \end{bmatrix}, \tag{104}$$

we arrive at

$$|\mathcal{R}_{uv}^{\alpha\beta}| \leq \|f\|_\infty \boldsymbol{\psi}_M{}^T \boldsymbol{R}_{uv}^{\alpha\beta} \boldsymbol{\psi}_M + \|f\|_\infty \int_{-1}^{1} \begin{bmatrix} U_k \\ V_k \end{bmatrix}^T \boldsymbol{\Sigma}_{uv}^{\alpha\beta} \begin{bmatrix} U_k \\ V_k \end{bmatrix} dx. \tag{105}$$

Recalling that $|\mathcal{L}_n(x)| \leq 1$ for all $n \geq 0$ [19], equation (35) follows from the estimate

$$\|f\|_\infty = \sup_{x\in[-1,1]} \left| \sum_{n=0}^{p} \hat{f}_n(\gamma)\mathcal{L}_n(x) \right| \leq \sum_{n=0}^{p} \left| \hat{f}_n(\gamma) \right| = \left\| \hat{\boldsymbol{f}}(\gamma) \right\|_1. \tag{106}$$

References

[1] R. P. Agarwal and D. O'Regan. *Ordinary and Partial Differential Equations, With Special Functions, Fourier Series and Boundary Value Problems*. Universitext. Springer-Verlag New York, 2009.

[2] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou. Input-Output Analysis of Distributed Parameter Systems Using Convex Optimization. In *IEEE 53rd Annu. Conf. Decis. Control (CDC), 2014*, pages 4310 – 4315, Los Angeles, CA, 2014. IEEE.

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

[4] E. D. Andersen, B. Jensen, J. Jensen, and R. Sandvik. MOSEK version 6 . MOSEK Technical report : TR-2009-3. Technical report, MOSEK ApS, Fruebjergvej 3 Box 16, 2100 , Copenhagen, Denmark, Copenhagen, 2009.

[5] D. Bertsimas and C. Caramanis. Geometry, Moments, and Semidefinite Optimization. In *Proc. 15th Int. Symp. MTNS*, pages 1–15, 2002.

[6] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[7] P. Constantin and C. R. Doering. Variational bounds in dissipative systems. *Phys. D Nonlinear Phenom.*, 82(3):221–228, 1995.

[8] P. Constantin and C. R. Doering. Variational bounds on energy dissipation in incompressible flows. II. Channel flow. *Phys. Rev. E*, 51(4):3192–3198, 1995.

[9] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience Publisher Inc., New York, 1st edition, 1953.

[10] C. R. Doering and P. Constantin. Energy dissipation in shear driven turbulence. *Phys. Rev. Lett.*, 69(11):1648–1651, 1992.

[11] C. R. Doering and P. Constantin. Variational bounds on energy dissipation in incompressible flows: Shear flow. *Phys. Rev. E*, 49(5):4087–4099, 1994.

[12] C. R. Doering and P. Constantin. Variational bounds on energy dissipation in incompressible flows. III. Convection. *Phys. Rev. E*, 53(6):5957–5981, may 1996.

[13] C. R. Doering and J. M. Hyman. Energy stability bounds on convective heat transport: Numerical study. *Phys. Rev. E*, 55(6):7775–7778, 1997.

[14] J. Dougall. The product of two Legendre polynomials. *Proc. Glas. Math. Assoc.*, 1(3):121–125, 1953.

[15] G. Fantuzzi and A. Wynn. Construction of an optimal background profile for the Kuramoto–Sivashinsky equation using semidefinite programming. *Phys. Lett. A*, 379(1-2):23–32, jan 2015.

[16] G. Fantuzzi and A. Wynn. Optimal bounds with semidefinite programming: An application to stress driven shear flows. *Phys. Rev. E*, 93(4):043308, 2016.

[17] M. Giaquinta and S. Hildebrandt. *Calculus of Variations I*, volume 310 of *Grundlehren der mathematischen Wissenschaften*. Springer Berlin Heidelberg, 1996.

[18] G. I. Hagstrom and C. R. Doering. Bounds on Surface Stress-Driven Shear Flow. *J. Nonlinear Sci.*, 24(1):185–199, 2014.

[19] D. Jackson. *The Theory of Approximation*, volume 11 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, New York, 1930.

[20] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *IEEE Int. Symp. Comput. Aided Control Syst. Des.*, pages 284 – 289, Taipei, 2004.

[21] J. Löfberg. Pre-and post-processing sum-of-squares programs in practice. *IEEE Trans. Automat. Contr.*, 54(5):1007–1011, 2009.

[22] A. Papachristodoulou and M. M. Peet. On the Analysis of Systems Described by Classes of Partial Differential Equations. *Proc. 45th IEEE Conf. Decis. Control*, pages 747–752, 2006.

[23] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Program. Ser. B*, 96(2):293–320, 2003.

[24] M. M. Peet and P.-A. Bliman. An Extension of the Weierstrass Theorem to Linear Varieties: Application to Delayed Systems. In *7th IFAC Work. Time-Delay Syst.*, pages 1–4, Nantes, France, 2007.

[25] B. Straughan. *The Energy Method, Stability, and Nonlinear Convection*, volume 91 of *Applied Mathematical Sciences*. Springer-Verlag New York, 2 edition, 2004.

[26] W. Tan. *Nonlinear Control Analysis and Synthesis using Sum-of-Squares Programming*. Phd thesis, University of california, Berkeley, 2006.

[27] W. Tang, C. P. Caulfield, and W. R. Young. Bounds on dissipation in stress-driven flow. *J. Fluid Mech.*, 510:333–352, 2004.

[28] G. Valmorbida, M. Ahmadi, and A. Papachristodoulou. Semi-definite programming and functional inequalities for Distributed Parameter Systems. In *53rd IEEE Conf. Decis. Control*, pages 4304 – 4309, Los Angeles, CA, 2014.

[29] Giorgio Valmorbida, Mohamadreza Ahmadi, and Antonis Papachristodoulou. Stability Analysis for a Class of Partial Differential Equations via Semidefinite Programming. *IEEE Trans. Automat. Contr.*, 61(6):1649–1654, 2016.

[30] Giorgio Valmorbida and Antonis Papachristodoulou. Introducing INTSOSTOOLS : A SOSTOOLS plug-in for integral inequalities. In *Proc. 2015 Eur. Control Conf.*, pages 1231–1236, Linz, Austria, 2015.

[31] 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*, 377(41):2931–2938, 2013.

[32] 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*, 92(4):043012, 2015.

[33] E. Zeidler. *Applied Functional Analysis - Applications to Mathematical Physics*, volume 108 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1st edition, 1995.