Math. Program., Ser. A (2006)

Digital Object Identifier (DOI) 10.1007/s10107-006-0702-z

Dimitris Bertsimas* · Constantine Caramanis**

Bounds on linear PDEs via semidefinite optimization

Received: September 19, 2001 / Accepted: December 14, 2005
Published online: 6 March 2006 – © Springer-Verlag 2006

Abstract. Using recent progress on moment problems, and their connections with semidefinite optimization, we present in this paper a new methodology based on semidefinite optimization, to obtain a hierarchy of upper and lower bounds on linear functionals defined on solutions of linear partial differential equations. We apply the proposed method to examples of PDEs in one and two dimensions, with very encouraging results. We pay particular attention to a PDE with oblique derivative conditions, commonly arising in queueing theory. We also provide computational evidence that the semidefinite constraints are critically important in improving the quality of the bounds, that is, without them the bounds are weak.

1. Introduction

In many real-world applications of phenomena that are described by partial differential equations (PDEs) we are primarily interested in a functional of the solution of the PDE, as opposed to the solution itself. For example, we might be interested in the average temperature along part of the physical boundary, rather than the entire distribution of temperature in a mechanical device; or we might be interested in the average inventory and its variability in a supply chain network; or finally, we might be interested in the expected queue lengths at various stations in a queueing network.

Given that analytical solutions of PDEs are scarce, there is a large body of literature on numerical methods for solving PDEs. Excellent references can be found in Quarteroni and Valli [19], Strang and Fix [23], Brezzi and Fortin [7]. Such methods typically involve some discretization of the domain of the solution, and thus obtain an approximate solution by solving the resulting equations, and matching boundary values and initial conditions. Such approaches scale exponentially with the dimension, i.e., if we use $O(1/\epsilon)$ points in each dimension, the size of systems we need to solve is of the order of $(1/\epsilon)^d$ for d–dimensional PDEs and result in accuracy of $O(\epsilon)$. There have been some efforts to control this explosion. An approach based on Lagrangean duality that performs computations using a coarse discretization, but provides bounds on the fine discretization solution based on the coarse discretization, is presented in Peraire and Patera

D. Bertsimas: Boeing Professor of Operations Research, Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, E53-363, Cambridge, MA 02142, USA.
e-mail: dbertsim@mit.edu

C. Caramanis: Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA.
e-mail: cmcaram@mit.edu.

* corresponding author. Research supported by the SMA-MI Talliance.

** Research partially supported by the SMA-MIT alliance and an NSF predoctoral fellowship.

[15], Paraschivoiu, Peraire and Patera [16] and Peraire and Patera [18]. For other duality based methods see Brezzi and Fortin [7]. These methods all provide approximations to the PDE and its linear functionals.

Given the interest in a functional of the solution of the underlying PDE, and the computational difficulty of obtaining close approximations, it is desirable to obtain bounds on the functional at a decreased computational burden. The discretization-based approaches yield solutions that are approximations to the PDE solution, and in particular, cannot provide guaranteed bounds on a functional of the solution.

The main focus of this work is to address exactly this point: providing guaranteed upper and lower bounds, with decreased computational burden.

Contributions

Using recent progress on moment problems and their connections with semidefinite optimization, we present a new methodology based on semidefinite optimization, to obtain a hierarchy of upper and lower bounds on linear functionals, as well as bounds on the supremum and infimum functionals, for linear PDEs with coefficients that are polynomials of the variables. The obtained bounds are guaranteed upper and lower bounds, and not simply approximations. Indeed, for this reason, even the results obtained with minimal computational effort may be of interest; while the upper and lower bounds may be loose initially, they are nevertheless guaranteed bounds, and hence contain potentially useful information. This is not the case for discretization-based solutions. If the discretization is too coarse, then the approximation may be unable to capture the essential features of the PDE, and the resulting solution may be useless.

We apply the proposed method to four examples of PDEs in one and two dimensions, with very encouraging results. The first three examples are quite standard examples from the literature, satisfying von Neumann, and Dirichlet boundary conditions. The last example we consider is a PDE with oblique derivative boundary conditions. This is motivated by a particular problem in queueing theory of networks. We discuss both the physical application and the solution more extensively.

In our discussion of these examples, we also provide computational evidence that the semidefinite constraints are critically important in improving the quality of the bounds, that is, without them the bounds are weak. The numerical results further indicate fast convergence. We pay particular attention to the quality of the bounds obtained with minimal computational effort, that is, the first few points on the convergence curve. The practicality and numerical stability of the proposed method depend on the numerical stability of semidefinite optimization codes, which are currently under intensive research. In addition, the stability of the formulation itself is important. In the formulation we propose, we consider semidefinite matrices of a particular form. It is possible that moment matrices lead to ill-conditioned problems, and thus that the formulation itself contributes to the instability of the numerical solutions. Future research should consider different parameterizations that lead to solutions with better stability properties. Specialized semidefinite solvers could also alleviate this problem, since the semidefinite constraints are all of the same form. We hope that progress in semidefinite optimization codes will lead to improved performance for obtaining bounds on PDEs using the method of the present paper.

Moment Problems and Semidefinite Optimization

Semidefinite optimization is currently at the center of much research activity in the area of mathematical programming, both from the point of view of new application areas (see for example the survey paper of Vandenberghe and Boyd [26]) as well as algorithmic development.

Problems involving moments of random variables arise naturally in many areas of mathematics, economics, and operations research. Recently, semidefinite optimization methods using moments have been applied to several problems arising in probability theory, finance and stochastic optimization. Bertsimas [2] applies semidefinite optimization methods to find bounds for stochastic optimization problems arising in queueing networks. Bertsimas and Popescu [4] and Lasserre [14] apply semidefinite optimization methods to find best possible bounds on the probability that a multi-dimensional random variable belongs in a set given a collection of its moments. In [5], Bertsimas and Popescu use these methods to find best possible bounds for pricing financial derivatives without assuming particular price dynamics. Lasserre [13] and Parrilo [17] present a method for global optimization of polynomials based on moments and semidefinite optimization.

Structure of the Paper

The paper is structured as follows. In Section 2, we present the proposed approach. In Section 3, we present four examples that show how the method works, how it performs numerically, and also illustrate the method's applicability to problems arising from concrete application areas. Finally, in Section 4, we provide some concluding remarks.

2. The Proposed Method

Suppose we are given partial differential operators L and G operating on some distribution space $\mathcal{A}$:

$$L, G : \mathcal{A} \longrightarrow \mathcal{A},$$

and we are interested in computing

$$\int Gu(\boldsymbol{x}),$$

where $u \in \mathcal{A}$ (note also that $f \in \mathcal{A}$) satisfies the PDE,

$$Lu(\boldsymbol{x}) = f(\boldsymbol{x}), \qquad \boldsymbol{x} = (x_1, \dots, x_d) \in \Omega \subset \mathbb{R}^d, \tag{1}$$

including the appropriate boundary conditions on $\partial\Omega$.

Eq. (1) is understood in the sense that both sides of the equation act in the same way on a given class of functions $\mathcal{D}$, i.e.,

$$Lu = f \iff \int (Lu)\phi = \int f\phi, \quad \forall \phi \in \mathcal{D},$$

where $\mathcal{D}$ is taken to be some sufficiently nice class of test functions—typically a subset of the smooth functions $\mathcal{C}^\infty$.

We assume that the operators L and G are linear operators of orders r_1, r_2 respectively, with coefficients that are polynomials of the variables. In Section 2.5, we discuss extensions to the max and min functionals, which are nonlinear. For the linear case, we have:

$$Lu(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|\le r_1} L_{\boldsymbol{\alpha}}(\boldsymbol{x})\frac{\partial^{\boldsymbol{\alpha}} u(\boldsymbol{x})}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}, \qquad Gu(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|\le r_2} G_{\boldsymbol{\alpha}}(\boldsymbol{x})\frac{\partial^{\boldsymbol{\alpha}} u(\boldsymbol{x})}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}},$$

where $\boldsymbol{\alpha} = (i_1, \dots, i_d)$ is a multi-index,

$$\frac{\partial^{\boldsymbol{\alpha}} u(\boldsymbol{x})}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}} = \frac{\partial^{\sum_k i_k} u(\boldsymbol{x})}{\partial x_1^{i_1}\cdots \partial x_d^{i_d}},$$

and $L_{\boldsymbol{\alpha}}(\boldsymbol{x})$ and $G_{\boldsymbol{\alpha}}(\boldsymbol{x})$ are multivariate polynomials (we discuss extensions in Section 2.6). We restrict ourselves to the case where $\mathcal{D}$ is separable, that is, it has a countable dense subset. In the context of the problems we consider, this restriction is rather mild. For instance, if u has compact support, then we can assume the elements of $\mathcal{D}$ have compact support, in which case by the Stone-Weierstrass theorem, $\mathcal{D}$ is separable. The condition that u have compact support may also be replaced by the (slightly) weaker condition that u have exponentially decaying tails.

Let $\mathcal{F} = \{\phi_1, \phi_2, \dots\}$ generate (in the basis sense) a dense subset of $\mathcal{D}$. Then, by the linearity of integration we have

$$\begin{aligned} Lu = f \iff \int (Lu)\phi &= \int f\phi, \quad \forall \phi \in \mathcal{D}, \\ \iff \int (Lu)\phi_i &= \int f\phi_i, \quad \forall \phi_i \in \mathcal{F}. \end{aligned}$$

We discuss different choices for the subset $\mathcal{F}$ in Section 2.6. This paper focuses on the subspace spanned by the monomials $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{i_1}\dots x_d^{i_d}$. Polynomials have the property that they are closed under action by polynomial coefficient differential operators.

We obtain bounds by imposing the equalities above for increasing finite subsets of $\mathcal{F}$. When the basis elements are monomials, we obtain upper and lower bounds on $\int Gu(\boldsymbol{x})$ by solving the optimization problem

$$\begin{aligned} &\text{max/min} : \textstyle\int Gu(\boldsymbol{x}) \\ &\text{such that} : \textstyle\int Lu(\boldsymbol{x})\boldsymbol{x}^{\boldsymbol{\alpha}} = \int f(\boldsymbol{x})\boldsymbol{x}^{\boldsymbol{\alpha}}, \quad \text{for all } \boldsymbol{x}^{\boldsymbol{\alpha}} \text{ with degree less than } N. \end{aligned}$$

We show that we can reformulate this problem as a semidefinite optimization problem. As such, it can be solved by black-box semidefinite solvers, such as the ones we use and discuss below.

The Adjoint Operator

The adjoint operator, L^*, is defined by the equation:

$$\int (Lu)\phi = \int u(L^*\phi), \qquad \forall u \in \mathcal{A}, \quad \forall \phi \in \mathcal{D}.$$

Therefore, if we have both L and L^*, then equality in the original PDE becomes:

$$\begin{aligned} Lu = f \iff & \int (Lu)\phi = \int f\phi, \quad \forall \phi \in \mathcal{D}, \\ \iff & \int (Lu)\phi_i = \int f\phi_i, \quad \forall \phi_i \in \mathcal{F}, \\ \iff & \int u(L^*\phi_i) = \int f\phi_i, \quad \forall \phi_i \in \mathcal{F}. \end{aligned} \tag{2}$$

To illustrate the computation of the adjoint operator, we consider the one–dimensional case. The general term of this operator is, up to a constant multiple:

$$x^a \frac{\partial^b}{\partial x^b}.$$

Using the notation $\tilde{\phi} = x^a\phi$, this term's contribution to the adjoint operator is as follows.

$$\begin{aligned} \int_\Omega x^a(\partial^b u)\phi = \int_\Omega (\partial^b u)(x^a\phi)\,dx = \int_\Omega (\partial^b u)\tilde{\phi}\,dx \\ = u^{(b-1)}\tilde{\phi}\Big|_{\partial\Omega} + \cdots + (-1)^{k+1}u^{(b-k)}\tilde{\phi}^{(k-1)}\Big|_{\partial\Omega} + \cdots \\ +(-1)^{b+1}u\tilde{\phi}^{(b-1)}\Big|_{\partial\Omega} + (-1)^b \int_\Omega u\partial^b\tilde{\phi}\,dx. \end{aligned}$$

The adjoint of a linear operator is composed of terms such as these. Thus, while perhaps notationally tedious in higher dimensions, computing the adjoint of a linear partial differential operator with polynomial coefficients is essentially only as difficult as performing the chain rule for differentiation on polynomials, and in particular, it may be easily automated.

2.1. Linear Constraints

We define variables in an optimization sense, that we will subsequently seek to constrain, and then optimize. We define variables corresponding to the full moments,

$$m_{\boldsymbol{\alpha}} = \int_\Omega \boldsymbol{x}^{\boldsymbol{\alpha}} u(\boldsymbol{x}) = \int_\Omega {x_1}^{i_1} \cdots {x_d}^{i_d} u(\boldsymbol{x}),$$

together with variables related to the boundary. These represent the integral of a monomial against the solution to the PDE, u, or some directional derivative of u, along some portion of the boundary, $\partial\Omega$. For instance, we could have:

$$z_{\boldsymbol{\alpha}} = \int_{\partial\Omega} \boldsymbol{x}^{\boldsymbol{\alpha}} u(\boldsymbol{x}) = \int_{\partial\Omega} {x_1}^{i_1} \cdots {x_d}^{i_d} u(\boldsymbol{x}).$$

The specific form of these variables depends on the nature of the boundary conditions we are given (see Section 3 for specific examples). Therefore, for notational convenience, we simply use $z = \{z_{\boldsymbol{\alpha}}\}$ to refer to the entire family of possible boundary moments. We refer to the quantities $m_{\boldsymbol{\alpha}}$ and $z_{\boldsymbol{\alpha}}$ as moments, even though $u(\cdot)$ may not be a probability distribution. Now consider Eqs. (2). We select as ϕ_i's the family of monomials $\{\boldsymbol{x}^{\boldsymbol{\alpha}}\}$. For the cases we consider, L, and thus L^*, are linear operators with coefficients that are polynomials in $\boldsymbol{x}$. Then, Eqs. (2) can be written as linear equations in terms of the variables $\boldsymbol{M} = \{m_{\boldsymbol{\alpha}}\}$ and $z = \{z_{\boldsymbol{\alpha}}\}$. Thus, each monomial test function generates a linear equation in $\boldsymbol{M}$ and z.

The general idea of using the adjoint equation to obtain relations that the moments must satisfy has appeared in [3], and [12]. A more systematic use of the adjoint equation appeared in a restricted context in [21], for studying the steady–state distribution of certain Markov processes. Semidefinite constraints are not used in [3], [12], or [21]. As we demonstrate below, the power of the semidefinite constraints is crucial for the performance of the method in this more general context.

2.2. *Objective Function Value*

In the cases we consider, the operator G is also a linear operator with coefficients that are polynomials of the variables (we consider an extension in Section 2.5). Then, the functional $\int Gu$ can also be expressed as a linear function of the variables $\boldsymbol{M}$ and z, again by obtaining the adjoint operator G^*. By minimizing or maximizing this particular linear function, we obtain upper and lower bounds on the value of the functional.

2.3. *Semidefinite Constraints*

Let us assume that the solution to the PDE is bounded from below, that is, $u(\boldsymbol{x}) \geq u_0$. The constant u_0 may be unknown. In certain cases, u_0 is naturally known; for example if $u(\boldsymbol{x})$ is a probability distribution, or if $u(\boldsymbol{x})$ represents temperature, then $u(\boldsymbol{x}) \geq 0$. In such cases, the formulation we present is greatly strengthened by the semidefinite constraints.

We consider the vectors $\boldsymbol{F}_n(\boldsymbol{x}) = [\boldsymbol{x}^{\boldsymbol{\alpha}}]_{|\boldsymbol{\alpha}|\leq n}$ and the semidefinite matrix $\boldsymbol{F}_n(\boldsymbol{x})\boldsymbol{F}_n(\boldsymbol{x})'$. Then the matrices

$$\int_{\Omega} (u(\boldsymbol{x}) - u_0)\boldsymbol{F}_n(\boldsymbol{x})\boldsymbol{F}_n(\boldsymbol{x})', \qquad \int_{\partial\Omega} (u(\boldsymbol{x}) - u_0)\boldsymbol{F}_n(\boldsymbol{x})\boldsymbol{F}_n(\boldsymbol{x})',$$

are also positive semidefinite for all n. Replacing $\int (u(\boldsymbol{x}) - u_0)\boldsymbol{x}^{\boldsymbol{\alpha}}$ by $(m_{\boldsymbol{\alpha}} - u_0 \int \boldsymbol{x}^{\boldsymbol{\alpha}})$, and similarly for the boundary moments, this leads to semidefinite constraints affine in the variables $(\boldsymbol{M}, u_0)$ and (z, u_0).

Obtaining the complete set of constraints for these variables, is an extension to multiple dimensions of the classical moment problem (see Akhiezer [1]). The problem is to determine, given some sequence of numbers and a support set K, whether it is a valid moment sequence, that is to say, whether the numbers given are indeed the moments of a nonnegative function or distribution supported on K. In one dimension, if $K = \mathbb{R}$, then a given sequence of numbers $\{m_i\}$ is the set of moments of some nonnegative function $u(x)$ (i.e., $m_i = \int_{-\infty}^{+\infty} x^i u(x) dx$) if and only if the matrix

$$M_{2n} = \begin{pmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \vdots & & \ddots & \vdots \\ m_n & & & m_{2n} \end{pmatrix}$$

is positive semidefinite for every n. In the case where $u(x)$ must have support on $K = [0, \infty)$, we need to add the additional constraint that the matrix

$$M_{2n+1} = \begin{pmatrix} m_1 & m_2 & \cdots & m_{n+1} \\ m_2 & m_3 & \cdots & m_{n+2} \\ \vdots & & \ddots & \vdots \\ m_{n+1} & & & m_{2n+1} \end{pmatrix}$$

also be positive semidefinite for all n. For a partial moment sequence, say, $\{m_i\}_{i \leq N}$ for some N, the necessary and sufficient conditions for the $\{m_i\}$ to be true moments, are that there exist a *positive semidefinite extension* to the sequence, i.e., there exist numbers $\{\hat{m}_i\}_{i \geq N+1}$ such that the matrices M_{2n} and M_{2n+1} given above, are positive semidefinite for all n. This concept of a *semidefinite extension* condition is important in the sequel.

In multiple dimensions, it is generally unknown which are the exact necessary and sufficient conditions for $\boldsymbol{M} = \{m_{\boldsymbol{\alpha}}\}$ and $z = \{z_{\boldsymbol{\alpha}}\}$ to be a valid moment sequence, when we are working over a general domain. For a wide class of domains, however, the positivstellensatz (see [6]) and in particular Schmüdgen's distinguished representations (see [22]) find such conditions. We review this work briefly, and use it to derive the necessary and sufficient conditions for $\boldsymbol{M} = \{m_{\boldsymbol{\alpha}}\}$ and $z = \{z_{\boldsymbol{\alpha}}\}$ to be a valid moment sequence for the cases we consider.

An Operator Approach

Given a closed subset Ω of $\mathbb{R}^d$, a sequence of numbers $\boldsymbol{M} = \{m_{\boldsymbol{\alpha}}\}$ defines a valid moment sequence if there exists a (nonnegative regular bounded Borel) measure μ such that

$$m_{\boldsymbol{\alpha}} = \int_\Omega \boldsymbol{x}^{\boldsymbol{\alpha}} d\mu, \quad \forall \boldsymbol{\alpha}.$$

There are additional technical details, such as the precise definition of the spaces in which μ must be taken to live, and their respective topologies. Since this paper is primarily concerned with projections via linear functionals onto $\mathbb{R}$, where the topology and notions of convergence are the usual ones, we ignore the more technical details.

We define the linear operator

$$Hf = \int_\Omega f(\boldsymbol{x}) d\mu.$$

Given a measure μ, H is completely defined. A given moment sequence $\boldsymbol{M}$ completely specifies the restriction of H to all polynomial functions. If the sequence $\boldsymbol{M}$ is valid, then it is necessary that $Hf \geq 0$ whenever $f \geq 0$ on Ω. A classical theorem says that this is also sufficient:

Theorem 1 (Haviland [9]). *If $\Omega \subseteq \mathbb{R}^n$ is closed, then $\boldsymbol{M} = \{m_{\boldsymbol{\alpha}}\}$ defines a valid moment sequence if and only if the linear operator H is nonnegative on all polynomials that are nonnegative on Ω.*

Theorem 1 implies that the problem of finding necessary and sufficient conditions for $\boldsymbol{M} = \{m_{\boldsymbol{\alpha}}\}$ and $\boldsymbol{z} = \{z_{\boldsymbol{\alpha}}\}$ to be a moment sequence reduces to checking the nonnegativity of the image of polynomials that are nonnegative on Ω. In one dimension, any polynomial that is nonnegative may be written as a sum of squares of other polynomials. Since the square of a polynomial may be written as a quadratic form, the nonnegativity of the operator reduces to matrix semidefiniteness conditions. The Motzkin polynomial in $\mathbb{R}^3$,

$$P(x, y, z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2,$$

is an example that shows that in higher dimensions, nonnegative polynomials may not have sum of squares decompositions (see Reznick [20] for more details and historical background). We are concerned with nonnegativity over a particular domain, Ω. There has recently been much work concerning the representation of polynomials that are positive over a given domain. We refer the interested reader to Lasserre [13] and Parrilo [17] for further details of these representation theorems, as well as connections to semidefinite optimization. Here, we use a result of Schmüdgen [22] that gives a representation of all polynomials that are positive over a compact finitely generated semi-algebraic set Ω, as defined in the theorem below. This leads to necessary and sufficient conditions for a moment sequence to be valid on Ω.

Theorem 2 (Schmüdgen [22]). *Suppose $\Omega := \{\boldsymbol{x} \in \mathbb{R}^n \ : \ f_i(\boldsymbol{x}) \geq 0, \ 1 \leq i \leq r\}$ is closed and bounded, where $f_i(\boldsymbol{x})$ are polynomials. Then a polynomial $g(\boldsymbol{x})$ is positive on Ω if and only if it is expressible as a sum of terms of the form*

$$h_I^2(\boldsymbol{x}) \prod_{i \in I} f_i(\boldsymbol{x}),$$

for $I \subseteq \{1, \ldots, r\}$, possibly $I = \emptyset$, and h_I some polynomial.

Theorems 1 and 2 lead to the following result.

Theorem 3. *Given $\boldsymbol{M} = \{m_{\boldsymbol{\alpha}}\}$, there exists a function $u(\boldsymbol{x}) \geq u_0$ such that*

$$m_{\boldsymbol{\alpha}} = \int_{\Omega} (u(\boldsymbol{x}) - u_0)\boldsymbol{x}^{\boldsymbol{\alpha}}, \quad \text{for all multi-indices } \boldsymbol{\alpha},$$

for a closed and bounded domain Ω of the form

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^d \ : \ f_1(\boldsymbol{x}) \geq 0, \ldots, f_r(\boldsymbol{x}) \geq 0\},$$

if and only if for all subsets $I \subseteq \{1, \ldots, r\}$, and all $n \geq 0$, the matrix obtained from the expression

$$\int_{\Omega} (u(\boldsymbol{x}) - u_0) \boldsymbol{F}_n(\boldsymbol{x}) \boldsymbol{F}_n(\boldsymbol{x})' \prod_{i \in I} f_i(\boldsymbol{x}), \tag{3}$$

by replacing $\int (u(\boldsymbol{x}) - u_0)\boldsymbol{x}^{\boldsymbol{\alpha}}$ by $m_{\boldsymbol{\alpha}}$, is positive semidefinite.

Proof. Let H denote the operator whose restriction to polynomial functions is defined by the moment sequence $\boldsymbol{M}$. If the moments $\boldsymbol{M}$ satisfy all the semidefinite conditions for some n, then $Hf \geq 0$ for every polynomial function f nonnegative on Ω, that is expressible as a sum of terms of the form of Theorem 2, with degree at most $2n$. By Theorem 2, every positive polynomial has such an expression for some n, and by Theorem 1, nonnegativity of all polynomials is sufficient for $\boldsymbol{M}$ to be a valid moment sequence. □

Examples of domains for which the above result applies include the unit ball in $\mathbb{R}^d$, which can be written as

$$B = \{\boldsymbol{x} \in \mathbb{R}^d \,:\, 1 - x_1^2 - \cdots - x_d^2 \geq 0\},$$

and the unit hypercube,

$$C = \{\boldsymbol{x} \in \mathbb{R}^d \,:\, x_i \geq 0, 1 - x_i \geq 0,\ 1 \leq i \leq d\},$$

as well as the nonconvex and disconnected domain:

$$D = \{0, 1\}^d = \{\boldsymbol{x} \in \mathbb{R}^d \,:\, x_i(x_i - 1) \geq 0, x_i(x_i - 1) \leq 0, 1 \leq i \leq d\}.$$

We next make the connection to semidefinite constraints explicit. While all the results can be easily generalized to d dimensions, for notational simplicity we consider $d = 2$, assume that $u_0 = 0$ and use Ω as the unit hypercube C in two dimensions. Note that in this case there are four functions,

$$f_1(x_1, x_2) = x_1,\ \ f_2(x_1, x_2) = 1 - x_1,\ \ f_3(x_1, x_2) = x_2,\ \ f_4(x_1, x_2) = 1 - x_2,$$

defining the set Ω. Thus, there are $2^4 = 16$ possible subsets I of $\{1, 2, 3, 4\}$. Each of these subsets gives rise to a particular sequence of semidefinite constraints as follows. Denoting the moment sequence as $\{m_{i,j}\}$, if it is a valid moment sequence, then for $I = \emptyset$ we must have

$$\begin{pmatrix} m_{0,0} & m_{1,0} & m_{0,1} & m_{2,0} & m_{1,1} & m_{0,2} & \cdots \\ m_{1,0} & m_{2,0} & m_{1,1} & m_{3,0} & m_{2,1} & m_{1,2} & \cdots \\ m_{0,1} & m_{1,1} & m_{0,1} & m_{2,1} & m_{1,2} & m_{0,3} & \cdots \\ m_{2,0} & m_{3,0} & m_{2,1} & m_{4,0} & m_{3,1} & m_{2,2} & \cdots \\ m_{1,1} & m_{2,1} & m_{1,2} & m_{3,1} & m_{2,2} & m_{1,3} & \cdots \\ m_{0,2} & m_{1,2} & m_{0,3} & m_{2,2} & m_{1,3} & m_{0,4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \succeq \mathbf{0}.$$

For $I = \{2\}$, we obtain

$$\begin{pmatrix} m_{0,0} - m_{1,0} & m_{1,0} - m_{2,0} & m_{0,1} - m_{1,1} & m_{2,0} - m_{3,0} & \cdots \\ m_{1,0} - m_{2,0} & m_{2,0} - m_{3,0} & m_{1,1} - m_{2,1} & m_{3,0} - m_{4,0} & \cdots \\ m_{0,1} - m_{1,1} & m_{1,1} - m_{2,1} & m_{0,1} - m_{1,1} & m_{2,1} - m_{3,1} & \cdots \\ m_{1,1} - m_{2,1} & m_{2,1} - m_{3,1} & m_{1,2} - m_{2,2} & m_{4,0} - m_{5,0} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \succeq \mathbf{0}.$$

In practice, of course, we can only enforce the semidefiniteness of the truncated matrices, say, those involving moments of total degree at most N. Proceeding in this way, we obtain 16 semidefinite constraints for each truncation level N. If Ω is the unit ball in d dimensions, we have exactly two semidefinite constraints for each N. As we allow N to grow, we get increasingly tighter necessary conditions. As in the one–dimensional case, given some partial set of moments, $\{m_{\boldsymbol{\alpha}}\}_{|\alpha|\leq N}$, we can impose extension semidefinite constraints: $\{m_{\boldsymbol{\alpha}}\}_{|\alpha|\leq N}$ is valid only if for every K, there exist numbers $\{\hat{m}_{\boldsymbol{\alpha}}\}_{N<|\alpha|\leq N+K}$ such that the truncated matrices of appropriate size are positive semidefinite. In the two–dimensional examples to follow, we involve the extension parameter K explicitly, in order to obtain direct numerical evidence of the power of the semidefinite constraints.

2.4. *The Overall Formulation*

We wish to solve for certain linear functionals of the PDE

$$Lu = f.$$

The variables of the optimization problem we formulate are the full moments $m_{\boldsymbol{\alpha}} = \int_{\Omega} \boldsymbol{x}^{\boldsymbol{\alpha}} u(\boldsymbol{x})$, and the boundary moments that may arise from integration against u or some directional derivative of u along some portion of the boundary, for example, $z_{\boldsymbol{\alpha}} = \int_{\partial\Omega} \boldsymbol{x}^{\boldsymbol{\alpha}} u(\boldsymbol{x})$. Also, we introduce the variable u_0, a lower bound on $u(\boldsymbol{x}), \boldsymbol{x} \in \Omega$. This value might be naturally known. The semidefinite optimization consists of linear equality constraints, and semidefinite constraints. The linear constraints are generated from the adjoint equation,

$$\int_{\Omega} u(L^*\phi) = \int_{\Omega} f\phi,$$

for different test functions ϕ. We focus primarily on monomial test functions $\phi_{\boldsymbol{\alpha}} = \boldsymbol{x}^{\boldsymbol{\alpha}}$. The semidefinite constraints express the fact that the variables $m_{\boldsymbol{\alpha}}$ and $z_{\boldsymbol{\alpha}}$ are in fact moments. Subject to these constraints, we maximize and minimize a linear function of the variables that expresses the given linear functional. The overall steps of the formulation process are then summarized as follows:

1. Compute the adjoint operator L^*.
2. Generate the i^{th} equality constraint, $1 \leq i \leq n$, by requiring that
$$\int u(L^*\phi_i) = \int f\phi_i.$$
This step is controlled by the degree–bound parameter N, which describes how many monomial test functions will be used, i.e., N determines the value of n above.
3. Generate the semidefinite constraints among the moments that appear in the linear constraints. Also, generate the semidefinite extension constraints. These are extensions of the original semidefinite constraints, and they include variables not present in the linear constraints. The level of extension is governed by the parameter K. Note that the semidefinite constraints only depend on the domain Ω and not on the operator L.
4. Compute upper and lower bounds on the given functional by solving a semidefinite optimization problem over the intersection of the positive semidefinite cone and the equality constraints.

For each pair (N, K), this algorithm produces upper and lower bounds to the value of the functional, and not just approximations. As these parameters increase, the upper and lower bounds converge to each other, thus, in principle producing exact values for the linear functional.

2.5. *The Maximum and Minimum Operator*

In some problems, such as some of those considered in the sequel, we naturally know the lower bound u_0, or perhaps we have a lower bound on u_0. In these cases, the semidefinite constraints strengthen the formulation significantly. On the other hand, if u_0 is not known, and if the linear constraints are strong, in the sense that they provide good constraints on the set of feasible moment vectors, then the semidefinite constraints can be used to obtain bounds on u_0. In particular, as we illustrate in Examples 1 and 2 below, we can use the semidefinite constraints to compute bounds on the (nonlinear) functionals max and min. Indeed, suppose that the given functional is $Gu = \min_{\boldsymbol{x} \in \Omega} u(\boldsymbol{x})$. Recall that we have defined a variable u_0, such that $u(\boldsymbol{x}) \geq u_0$ for all $\boldsymbol{x}$ in the domain Ω. Thus the function $\hat{u} := (u - u_0)$ satisfies $\hat{u}(\boldsymbol{x}) \geq 0$, for all $\boldsymbol{x}$ in Ω. Therefore the corresponding moments, $\{\hat{m}_{\boldsymbol{\alpha}}\}$, computable from the moments of $u(\boldsymbol{x})$ by the relation

$$\hat{m}_{\boldsymbol{\alpha}} = m_{\boldsymbol{\alpha}} - u_0 \int_{\Omega} \boldsymbol{x}^{\boldsymbol{\alpha}},$$

satisfy all the semidefinite constraints, for every truncation $|\alpha| \leq N$. Conversely, if $u(\boldsymbol{x}_0) - u_0 < 0$ for some $\boldsymbol{x}_0$ in Ω, then by the sufficiency of the full semidefinite conditions, there exists some N, and some semidefinite matrix condition, that at truncation N (and higher) is violated by the moments of $(u(\boldsymbol{x}) - u_0)$. Therefore, if for some chosen matrix truncation level N, we solve for $u_{0,N}$ the maximum value for which the N–truncated matrices are all positive semidefinite, then we obtain an upper bound $u_{0,N} \geq u_0$, such that $u_{0,N}$ converges to u_0 as $N \rightarrow \infty$. In our case, however, we do not have the actual moments $\{\hat{m}_{\boldsymbol{\alpha}}\}$ of u, but rather some bounds, $\underline{m}_{\boldsymbol{\alpha}} \leq m_{\boldsymbol{\alpha}} \leq \overline{m}_{\boldsymbol{\alpha}}$. Therefore, for any truncation N, the maximum value $\tilde{u}_{0,N}$ for which the N–truncated matrices of *some* sequence of numbers satisfying the moment bounds above, are all positive semidefinite, will be at least $u_{0,N}$. This follows because the true moments of u are of course contained in the above moment bounds, and hence by allowing ourselves to look at possibly other values within the bounds, the maximization can only yield a value at least as large. Thus we have shown the following.

Proposition 1. *If in the formulation of the semidefinite program as described above, we maximize the variable u_0, we obtain an upper bound to the actual minimum of $u(\boldsymbol{x})$ over its domain Ω.*

A lower bound on the true value of the maximum of $u(\boldsymbol{x})$ over its domain Ω, may similarly be obtained.

Trying to find a lower bound on the minimum (respectively an upper bound on the maximum) by this method, can only yield uninformative answers (i.e., $-\infty$), since for any $u_* \leq \min u(\boldsymbol{x})$, $u(\boldsymbol{x}) - u_*$ is always nonnegative.

Note that here the semidefinite constraints are critical, as without them we could not obtain any information about the minimum or maximum of the function $u(\boldsymbol{x})$ over its domain. This is because the additional variable u_0 is introduced linearly, and because of the linearity of integration, cannot be calculated by the family of linear constraints.

2.6. Using Trigonometric Moments

Instead of choosing polynomials as test functions, we could choose other classes of test functions. Polynomials are particularly convenient as they are closed under differentiation. While this property is not a necessary condition for the proposed method to work, it significantly limits the proliferation of variables we introduce. When the linear operator has coefficients that are not polynomials, other bases might be more appropriate.

The trigonometric functions $\{\sin(nx), \cos(nx)\}$ are also closed under differentiation (again we can form products in higher dimensions, just as with monomials). Using trigonometric functions as a basis of our test functions provides a straightforward way to deal with linear operators with trigonometric coefficients. Indeed, the choice of test function basis should depend on the coefficients of the linear operator. In Section 3.5, we present an example of the use of the method with trigonometric test functions.

3. Examples

In this section, we illustrate our approach with four examples: (1) a simple homogeneous ordinary differential equation, (2) a more interesting ODE: Bessel's equation, (3) a two–dimensional partial differential equation (PDE) known as Helmholtz's equation, and (4) a two–dimensional PDE describing a diffusion, with oblique derivative conditions. This is motivated by an application to network queueing theory. We discuss this more extensively in Section 3.4.

3.1. Example 1: $u'' + 3u' + 2u = 0$

We consider the linear ODE with constant coefficients

$$u'' + 3u' + 2u = 0, \tag{4}$$

with $\Omega = [0, 1]$ and with the boundary conditions $u'(0) = -2e^2$ and $u'(1) = -2$. In this case, we can easily analytically compute the solution $u(x) = e^2 \cdot e^{-2x}$. We apply the proposed method to obtain bounds on the moments. For simplicity of the exposition we use the fact that $u(x) \geq 0$.

We can compute the adjoint operator directly by integration by parts:

$$\int_0^1 (u'' + 3u' + 2u)\phi = u'\phi\Big|_0^1 - u\phi'\Big|_0^1 + 3u\phi\Big|_0^1 + \int_0^1 (u\phi'' - 3u\phi' + 2u\phi).$$

We use $\phi_i(x) = x^i$, $i = 0, \ldots, N$ and let

$$m_i = \int_0^1 x^i u(x)dx.$$

Together with the two unknown boundary conditions $u(0)$ and $u(1)$, we have $N+1$ variables $m_i, i = 0, \dots, N$ for a total of $N+3$ variables. The linear equality constraints generated by the adjoint equations are:

$$\begin{aligned}
\phi = 1 &: \Rightarrow 3(u(1) - u(0)) + 2m_0 = u'(0) - u'(1),\\
\phi = x &: \Rightarrow 2u(1) + u(0) - 3m_0 + 2m_1 = -u'(1),\\
\phi = x^2 &: \Rightarrow u(1) + 2m_0 - 6m_1 + 2m_2 = -u'(1),\\
&\ \ \vdots\\
\phi = x^N &: \Rightarrow (3-N)u(1) + N(N-1)m_{N-2} - 3N \cdot m_{N-1} + 2m_N = -u'(1).
\end{aligned}$$

Since we assume that the solution has support on [0, 1], we apply Proposition 3 to derive the two semidefinite constraints:

$$\begin{pmatrix} m_0 & m_1 & \cdots & m_N \\ m_1 & m_2 & \cdots & m_{N+1} \\ \vdots & & \ddots & \vdots \\ m_N & & & m_{2N} \end{pmatrix} \succeq \mathbf{0}, \qquad \begin{pmatrix} m_0 - m_1 & m_1 - m_2 & \cdots & m_N - m_{N+1} \\ m_1 - m_2 & m_2 - m_3 & \cdots & m_{N+1} - m_{N+2} \\ \vdots & & \ddots & \vdots \\ m_N - m_{N+1} & & & m_{2N} - m_{2N+1} \end{pmatrix} \succeq \mathbf{0}.$$

Subject to these constraints, we maximize and minimize each of the m_i, $0 \le i \le N$, in order to obtain upper and lower bounds for m_i.

We applied two semidefinite optimization packages to solve the resulting SDPs: the optimization package SDPA version 5.00 by Fujisawa, Kojima and Nakata [8] and the Matlab-based package SeDuMi version 1.03, by Sturm [24]. We ran the semidefinite optimizations on a Sparc 5.

In Table 1, we report the results from SDPA using monomials up to $N = 14$. As SDPA exhibited some numerical instability, we replaced the equality constraints $\boldsymbol{a}'\boldsymbol{x} = b$ with $-\varepsilon + b \le \boldsymbol{a}'\boldsymbol{x} \le b + \varepsilon$ with $\varepsilon = 0.001$. We see, as we would expect, that the performance begins to deteriorate as we ask for bounds on higher order moments.

In Table 2, we report results using SeDuMi with $N = 60$. SeDuMi successfully solved for the first 45 moments, such that the upper and lower bounds agree to 5 decimal points.

By implementing the method we outlined in Section 2.5 to compute the minimum and maximum of $u(x)$, and using SeDuMi, we obtain the exact value for the maximum of the function $u(x)$ over [0, 1], to be $u_0 = 7.389$. Again we note that it is the semidefinite constraints that allow us to compute upper (resp. lower) bounds on the nonlinear functional min (resp. max).

3.2. Example 2: The Bessel Equation

In this section, we consider Bessel's differential equation

$$x^2u'' + xu' + (x^2 - p^2)u = 0.$$

Table 1. Upper and lower bounds for the ODE (4) for $N = 14$, using SDPA. The total computation time was less than 15 seconds for all twelve SDPs

Variable	LB	UB
m_0	3.1939	3.1951
m_1	1.0969	1.1619
m_2	0.5970	0.5997
m_3	0.3957	0.3961
m_4	0.2916	0.3179
m_5	0.2280	0.8809

Table 2. Upper and lower bounds for the ODE (4) for $N = 60$, using SeDuMi. The total computation time was under five minutes. For this simple example, our method computes upper and lower bounds for the moments, that essentially coincide

Variable	LB	UB
m_0	3.1945	3.1945
m_1	1.0973	1.0973
m_2	0.5973	0.5973
m_3	0.3959	0.3959
m_4	0.2918	0.2918
m_5	0.2295	0.2295
m_6	0.1884	0.1884
m_7	0.1595	0.1595
m_8	0.1382	0.1382
m_9	0.1218	0.1218
m_{10}	0.1088	0.1088
$\vdots$	$\vdots$	$\vdots$
m_{20}	0.0524	0.0524
$\vdots$	$\vdots$	$\vdots$
m_{30}	0.0344	0.0344
$\vdots$	$\vdots$	$\vdots$
m_{40}	0.0256	0.0256

The Bessel function and its variants appear in one form or another in a wide array of engineering applications, and applied mathematics. Furthermore, while there are integral and series representations, the Bessel function is not expressible in closed form. The series representation of the Bessel function, which can be found in, e.g., Watson [28], is:

$$J_p(x) := \sum_{k=0}^{\infty} \frac{(-1)^k (x/2)^{2k+p}}{k!(k+p)!}.$$

Also, over the appropriate range, the Bessel function is neither nonnegative, nor convex.

In order to avoid numerical difficulties from large constant factors, we solve a modified version of Bessel's equation:

$$x^2 u'' + xu' + (49x^2 - p^2)u = 0. \tag{5}$$

The solution is $u(x) = J_p(7x)$. Assuming we are given the value of the derivatives on the boundary, using the monomials as the test functions, we obtain the adjoint equations:

$$\begin{aligned}
\phi = 1 &: \Rightarrow -u(1) + (1 - p^2)m_0 + 49m_2 = u'(1), \\
\phi = x &: \Rightarrow -2u(1) + (4 - p^2)m_1 + 49m_3 = u'(1), \\
\phi = x^2 &: \Rightarrow -3u(1) + (9 - p^2)m_2 + 49m_4 = u'(1), \\
&\quad \vdots \\
\phi = x^N &: \Rightarrow -(N+1)u(1) + ((N+1)^2 - p^2)m_N + 49m_{N+2} = u'(1).
\end{aligned}$$

In what follows, we choose $p = 1$. We used SeDuMi to compute the moments, and also to compute the max and min. Recall from the discussion in Section 2.5 that while we are able to obtain both upper and lower bounds for the moments, our method can only compute upper bounds to the minimum of the solution, and lower bounds for the maximum. Indeed, in the case of the Bessel function, the bounds we obtain for the minimum are greater than the actual value, and the bounds for the maximum are less than the actual value. The true values are: min $= -0.347$ and max $= 0.583$. In Table 3, we report the results from SeDuMi. SeDuMi reported severe numerical instabilities for the computation of the maximum for the cases $N = 30$ and $N = 40$.

Next, we use a loose lower bound of $u_0 = -0.4$, and we use SeDuMi to obtain bounds on the moments. We give the first few in Table 4. As in Example 1, the results are accurate (upper and lower bounds agree) to several decimal points.

3.3. *Example 3: The Helmholtz Equation*

In this section, we consider the two–dimensional PDE

$$\triangle u + k^2 u = f, \tag{6}$$

over $\Omega = [0, 1]^2$. We fix $k = 1$. Rather than a fixed number of boundary variables, in two dimensions we have $O(N)$ boundary variables. Partly for this reason, the semidefinite constraints prove particularly critical for providing interesting (non-trivial) bounds to the moments.

To compute the adjoint operator we need to use Stokes's theorem:

$$\int_\Omega d\omega = \int_{\partial\Omega} \omega.$$

We use the standard form of Green's theorem here:

$$\begin{aligned}
\int_\Omega \frac{\partial^2 u}{\partial x^2} \cdot \phi \, dxdy &= \int_\Omega u_{xx}\phi \, dxdy \\
&= -\int_\Omega u_x \phi_x \, dxdy + \int_{\partial\Omega} n_x u_x \phi \, dS \\
&= \int_\Omega u\phi_{xx} \, dxdy + \int_{\partial\Omega} (n_x u_x \phi - n_x u \phi_x) \, dS,
\end{aligned}$$

Table 3. Lower bounds for the maximum, and upper bounds for the minimum of the solution of Eq. (5) using SeDuMi

N	Minimum	Maximum
20	-0.3087	0.4986
24	-0.3101	0.5068
30	-0.3111	0.5081
40	-0.3142	0.5046

Table 4. Upper and lower bounds on the moments of the solution to Eq. (5) for $N = 24$, using SeDuMi

Variable	LB	UB
m_1	0.1766	0.1766
m_2	0.0903	0.0903
m_3	0.0583	0.0583
m_4	0.0438	0.0438
m_5	0.0361	0.0361

where the boundary $\partial\Omega$ of Ω is oriented to have unit outward normal (n_x, n_y). Thus, for Ω the unit square, we have,

$$\int_{\partial\Omega} n_x u_x \phi \, dS = \int_0^1 (u_x(x=1,y)\phi(x=1,y) - u_x(x=0,y)\phi(x=0,y))\, dy.$$

Again we consider the family of monomials,

$$\mathcal{F} = \{x^i \cdot y^j\}, \quad \text{for } i, j \in \mathbb{N} \cup \{0\}.$$

In addition to the variables

$$m_{i,j} = \int_0^1 \int_0^1 x^i y^j u(x,y) dx dy,$$

we also introduce the boundary moment variables denoting the integral of $u(x, y)$, or the partial derivatives $u_x = \frac{\partial}{\partial x}u$ and $u_y = \frac{\partial}{\partial y}u$, over some portion of the boundary:

$$\begin{aligned}
b_i^{x=1} &:= \int_0^1 u(x=1,y)y^i\, dy, & d_i^{x=1} &:= \int_0^1 u_x(x=1,y)y^i\, dy \\
b_i^{x=0} &:= \int_0^1 u(x=0,y)y^i\, dy, & d_i^{x=0} &:= \int_0^1 u_x(x=0,y)y^i\, dy \\
b_i^{y=1} &:= \int_0^1 u(x,y=1)x^i\, dx, & d_i^{y=1} &:= \int_0^1 u_y(x,y=1)x^i\, dx \\
b_i^{y=0} &:= \int_0^1 u(x,y=0)x^i\, dx, & d_i^{y=0} &:= \int_0^1 u_y(x,y=0)x^i\, dx.
\end{aligned}$$

We note again here that rather than the constant number of boundary variables in the one–dimensional case, we have $O(N)$ boundary variables. The adjoint relationship above yields:

$$
\begin{aligned}
\phi = 1 &: \Rightarrow d_0^{x=1} - d_0^{x=0} + d_0^{y=1} - d_0^{y=0} + m_{0,0} = \int_\Omega f\, dxdy \\
\phi = x^i &: \Rightarrow d_0^{x=1} + i(i-1)m_{i-2,0} + d_i^{y=1} - d_i^{y=0} + m_{i,0} \\
&= \int_\Omega f \cdot x^i\, dxdy + ib_0^{x=1} \\
\phi = y^j &: \Rightarrow d_j^{x=1} - d_j^{x=0} + d_0^{y=1} + j(j-1)m_{0,j-2} + m_{0,j} \\
&= \int_\Omega f \cdot y^i\, dxdy + jb_0^{y=1} \\
\phi = x^i y^j &: \Rightarrow d_j^{x=1} + i(i-1)m_{i-2,j} + d_i^{y=1} + j(j-1)m_{i,j-2} + m_{i,j} \\
&= \int_\Omega f \cdot x^i y^j\, dxdy + ib_j^{x=1} + jb_i^{y=1}.
\end{aligned}
$$

Either the $\{d_i^x, d_j^y\}$, or the $\{b_i^x, b_j^y\}$, are given as boundary values. In order to compare with the exact solution, we select the boundary conditions such that $u(x, y) = e^{x+y}$, and we assume we are given $\{b_i^x, b_j^y\}$.

Role of Semidefinite Constraints

In part because of the proliferation of variables due to the boundary variables, the semidefinite constraints are particularly important in higher dimensions for convergence of the upper and lower bounds. To highlight this claim in our numerical examples, we have introduced another parameter in the algorithm, relating specifically to the truncation level of the moment matrices which we demand to be positive semidefinite. We let N denote the degree in each dimension of the monomials used in the adjoint equation. For the Helmholtz equation, this results in linear equations using moments up to $x^N y^N$, where x, y are the two coordinates. Therefore the total degree of the highest order moment is $2N$, and the semidefinite matrices have rows and columns indexed by moments of degree at most N. As discussed and explained in Section 2.3, in addition to semidefinite conditions on some finite moment sequence, we may also impose semidefinite extension conditions, i.e., we may demand that numbers exist that can extend the given semidefinite matrix to a larger one that is again semidefinite. We introduce a parameter K that indexes the degree of the semidefinite extension constraints that the moments must satisfy. Thus the parameter pair (N, K) indicates that the moments are constrained by the linear equations generated by the monomials $x^i y^j$, for $0 \leq i, j \leq N$, and furthermore, they must satisfy the semidefinite moment matrix constraints, truncated to level $|\alpha| \leq 2(N + K)$.

For this example, the full moments are given by integration over the unit square. The boundary moments are given by integration over the unit interval. Therefore, as discussed in Section 2.3, the full moments must satisfy 16 semidefinite constraints for each truncation level, while the boundary moments must each satisfy two semidefinite constraints for each truncation level. The number of monomials used to obtain linear constraints is controlled by the parameter N, and the truncation level of the semidefinite matrices is governed by $(N + K)$.

The true results, as computed by Maple 6.0, are reported in Table 5. We ran the algorithm for various levels of N and K. First, using the commercial software AMPL, which uses CPLEX, we obtained upper and lower bounds using only the linear equations generated, and not imposing any semidefinite constraints (this may be thought of as the $K = -N$ case). The results without the semidefinite constraints are quite poor. They are given in Table 6.

Next, we ran the algorithm for much smaller values of N: $N = 3, 5, 6$, but included the accompanying semidefinite constraints. For $N = 3$, we used $K = 0$, and $K = 4$. The results are in Table 7. For $N = 5, 6$ we used $K = 0$, thus imposing semidefinite constraints truncated at level $|\alpha| \leq 2N$. These results are contained in Table 8. As the numbers illustrate, the improvement is significant. Furthermore, the fact that interesting bounds may apparently be obtained for small values of N and K, illustrates that the global nature of the moment variables seems appropriate for global performance measures, such as averages and variances.

3.4. *Example 4: Reflected Brownian Motion*

In this section, we consider a PDE with oblique derivative boundary conditions that describes the distribution of a diffusion approximation to the queue–length process of a two–station queue under heavy traffic conditions. In general, computing steady–state performance measures of network queueing systems can be difficult, and there is a wealth of literature addressing such problems, see for instance [3], [10], [25] and the references therein.

Consider two parallel finite–capacity queues, Q_1 and Q_2, coupled so that when Q_1 empties, it takes customers from Q_2, and if Q_2 is full, it transfers customers to Q_1. In the heavy traffic limit where the arrival rate approaches the service rate, the vector queue–length process can be approximated by a diffusion, or a reflected Brownian motion; see, for example, Harrison [10], Varadhan and Williams [27], and Harrison, Landau, and Shepp [11]. This is a diffusion, or reflected Brownian motion, inside the unit square. It satisfies Laplace's equation, $\triangle u = 0$, on the interior, and oblique derivative boundary conditions on the four edges, given by the four angles $\{\theta_1, \theta_2, \theta_3, \theta_4\}$. Thus along edge i, which we denote Γ_i, the derivative of the distribution of the diffusion in the direction θ_i

Table 5. Exact Results for the solution of the PDE (6)

Variable	Value
$m_{0,0}$	2.9525
$m_{1,0}$	1.7183
$m_{1,1}$	1.0000
$m_{2,0}$	1.2342
$m_{2,1}$	0.7183
$m_{2,2}$	0.5159
$m_{3,0}$	0.9681
$m_{3,1}$	0.5634
$m_{3,2}$	0.4047
$m_{3,3}$	0.3175

Table 6. Upper and lower bounds from linear optimization for $N = 5$, $N = 10$, $N = 20$. Significantly, even for $N = 20$, the bounds are quite loose, and some, for instance $m_{0,0}$, are completely uninformative

Variable	LB, $N = 5$	UB, $N = 5$	LB, $N = 10$	UB, $N = 10$	LB, $N = 20$	UB, $N = 20$
$m_{0,0}$	0.0000	$+\infty$	0.0000	$+\infty$	0.0000	$+\infty$
$m_{1,0}$	0.0000	4.3142	0.0000	4.0822	0.0000	3.8694
$m_{1,1}$	0.7559	1.0881	0.8557	1.0419	0.9400	1.0120
$m_{2,0}$	0.0000	4.6790	0.0000	4.6790	0.0000	4.6790
$m_{2,1}$	0.0545	1.0563	0.1417	1.0059	0.1749	0.9753
$m_{2,2}$	0.0000	0.9447	0.0000	0.9087	0.0000	0.9087
$m_{3,0}$	0.0000	4.8743	0.0000	4.4932	0.0000	4.4414
$m_{3,1}$	0.1692	0.6806	0.4015	0.6063	0.5025	0.5758
$m_{3,2}$	0.0000	0.6383	0.0000	0.6105	0.0000	0.5989
$m_{3,3}$	0.0000	0.5291	0.1684	0.3640	0.2573	0.3296

Table 7. Upper and lower bounds from the semidefinite optimization for $N = 3$, where we first use $K = 0$ and $K = 4$. The improvement can be entirely attributed to the strengthened semidefinite constraints, controlled by the parameter K. Note that already the bounds are quite good, in many cases revealing more than one significant digit, where as the bounds obtained without imposing the semidefinite constraints in Table 6 for much larger N, are significantly worse

Variable	LB, $N = 3, K = 0$	UB, $N = 3, K = 0$	LB, $N = 3, K = 4$	UB, $N = 3, K = 4$
$m_{0,0}$	2.5760	3.2009	2.9235	3.1707
$m_{1,0}$	1.5393	1.7742	1.6944	1.7742
$m_{1,1}$	0.9272	1.0428	0.9847	1.0130
$m_{2,0}$	0.9691	1.3657	1.1123	1.3088
$m_{2,1}$	0.5951	0.9491	0.7151	0.7458
$m_{2,2}$	0.3677	0.5765	0.4948	0.5456
$m_{3,0}$	0.7981	1.4707	0.8244	1.0478
$m_{3,1}$	0.3048	0.7928	0.5054	0.5818
$m_{3,2}$	0.0905	0.5717	0.3874	0.5133
$m_{3,3}$	0.0530	0.6945	0.3017	0.3399

with respect to the inward normal, vanishes. This corresponds to the Brownian motion being *reflected* instantaneously off of side k at an angle $-\theta_k$ from the normal. The x and y values of the diffusion represent the normalized number of customers in the first and second queue, respectively. Then, the problem we wish to solve is:

$$\begin{cases} \triangle u = 0 \\ \cos\theta_k \left(\frac{\partial u(z)}{\partial n}\right) + \sin\theta_k \left(\frac{\partial u(z)}{\partial \sigma}\right) = 0, \quad z \in \Gamma_k. \end{cases}$$

For the two–dimensional case, Trefethen and Williams [25] have obtained solutions to this class of problems by using the Schwarz–Christoffel transformations to find the real part of the analytic function (which, as they show, always exists) which conformally maps the domain of the diffusion to another region with straight lines, with orientation given by the angles of the oblique derivatives given in the problem definition. Because it relies on conformal mapping, their method applies strictly to the two–dimensional case.

The performance measures of interest are the expected value of the queue length at each of the two queues:

$$E_u(x) = \int_\Omega xu(x, y), \quad E_u(y) = \int_\Omega yu(x, y),$$

Table 8. Upper and lower bounds from the semidefinite optimization for $N = 5$, and $N = 6$, where we have $K = 0$ for both. Again we see improvement, in some cases significant, over the $N = 3$ case. Still, we remark that the most dramatic improvement is demonstrated in Table 7, when we first add the semidefinite extension constraints, from ($N = 3$, $K = 0$), to ($N = 3$, $K = 4$)

Variable	LB, $N = 5$, $K = 0$	UB, $N = 5$, $K = 0$	LB, $N = 6$, $K = 0$	UB, $N = 6$, $K = 0$
$m_{0,0}$	2.9346	2.9848	2.9493	2.9568
$m_{1,0}$	1.7122	1.7402	1.7122	1.7402
$m_{1,1}$	0.9966	1.0009	0.9972	1.0009
$m_{2,0}$	1.2299	1.2371	1.2310	1.2371
$m_{2,1}$	0.7147	0.7245	0.7156	0.7245
$m_{2,2}$	0.5113	0.5233	0.5132	0.5233
$m_{3,0}$	0.9639	0.9734	0.9646	0.9734
$m_{3,1}$	0.5605	0.5762	0.5608	0.5762
$m_{3,2}$	0.4003	0.4120	0.4021	0.4120
$m_{3,3}$	0.3106	0.3251	0.3146	0.3251

and also the integral with respect to arc length over each of the boundaries, Γ_k, representing the rate at which the diffusion hits Γ_k:

$$I_k = \int_{\Gamma_k} u(x, y).$$

Similarly to the previous examples considered, we obtain an adjoint relationship (see [25]):

$$\int_{\Omega} u \triangle \phi + \sum_{k=1}^{4} \int_{\Gamma_k} u \left(\frac{\partial}{\partial n} \phi - \tan \theta_k \frac{\partial}{\partial \sigma} \phi \right) = 0,$$

where as usual, ϕ is our test function, which we take to be a monomial. Here, $\frac{\partial}{\partial n}$ is the normal derivative, and $\frac{\partial}{\partial \sigma}$ is the tangent derivative, where the boundary is taken to be oriented counterclockwise, consistent with the inward pointing normal direction. The variables in this case are the full moments, $\{m_{ij}\}$, and the boundary moments, $\{b_i^{x=0}\}$, $\{b_i^{x=1}\}$, $\{b_i^{y=0}\}$, and $\{b_i^{y=1}\}$, where $I_1 = b_0^{x=0}$, $I_2 = b_0^{y=0}$, $I_3 = b_0^{x=1}$, and $I_4 = b_0^{y=1}$. Thus in this example, we are interested in bounding not only the full moments, m_{01} and m_{10}, but also the four boundary moments. Imposing semidefinite moment conditions on both the full moments, and also the boundary moments, we obtain the results given in Table 9.

3.5. *Trigonometric Test Functions*

In this section, we illustrate the use of trigonometric test functions. We consider the differential equation

$$u'' + 2u' + \sin(2\pi x)u = 10 \sin x - 20 \cos x + (10 - 10 \sin x) \sin(2\pi x). \qquad (7)$$

If we attempt to use a polynomial basis, we encounter a proliferation of variables, since the polynomials are not closed by action of the adjoint (which has a $\sin(2\pi x)$ term). We use the family of test functions

$$\phi_{2n}(x) := \sin(2\pi n x), \quad \phi_{2n+1}(x) := \cos(2\pi n x).$$

Table 9. Upper and lower bounds for performance measures of the queueing network example, for $N = 3, 6$, and 12, using SDPA. The actual values are: $I_1 = I_4 = 0.805295$, $I_2 = I_3 = 1.610589$, $m_{1,0} = 0.551506$, and $m_{0,1} = 0.448494$

Variable	LB, $N = 3$	UB, $N = 3$	LB, $N = 6$	UB, $N = 6$	LB, $N = 12$	UB, $N = 12$
$m_{1,0}$	0.5010	0.5917	0.5502	0.5532	0.5515	0.5516
$m_{0,1}$	0.4092	0.4988	0.4475	0.4509	0.4481	0.4488
I_1	0.5011	0.9953	0.7771	0.8328	0.8017	0.8086
$I_2 = I_3$	1.5007	1.6798	1.6077	1.6134	1.6104	1.6107
I_4	0.5011	0.9953	0.7771	0.8328	0.8017	0.8086

We define the variables:

$$
\begin{aligned}
m_{2n} &:= \int_\Omega u(x)\phi_{2n}(x)\,dx \\
m_{2n+1} &:= \int_\Omega u(x)\phi_{2n+1}(x)\,dx.
\end{aligned}
$$

The adjoint equations become:

$$
\begin{aligned}
\phi_1 = 1 : \Rightarrow 2u(1) - 2u(0) + m_2 &= \int_\Omega f\,dx + u'(0) - u'(1), \\
\phi_{2n} = \sin(2\pi nx) : \Rightarrow 2\pi n(u(0) - u(1)) + \frac{1}{2}m_{2(n-1)+1} - \frac{1}{2}m_{2(n+1)+1} & \\
-4\pi^2 n^2 m_{2n} - 4\pi n m_{2n+1} &= \int_\Omega f\phi_{2n}\,dx \\
\phi_{2n+1} = \cos(2\pi nx) : \Rightarrow 2u(1) - 2u(0) + \frac{1}{2}m_{2(n+1)} - \frac{1}{2}m_{2(n-1)} & \\
-4\pi^2 n^2 m_{2n+1} - 4\pi n m_{2n} &= \int_\Omega f\phi_{2n+1}\,dx + u'(0) - u'(1).
\end{aligned}
$$

We assume we are given $u'(0)$ and $u'(1)$. The products $\cos(2\pi nx) \cdot \sin(2\pi kx)$ appear in the semidefinite constraints. These can be rewritten as follows:

$$
\begin{aligned}
\sin(2\pi nx) \cdot \cos(2\pi kx) &= \frac{1}{2}(\sin(2\pi(n+k)x) + \operatorname{sgn}(n-k)\sin(2\pi|n-k|x)) \\
\sin(2\pi nx) \cdot \sin(2\pi kx) &= \frac{1}{2}(\cos(2\pi(n-k)x) - \cos(2\pi(n+k)x)) \\
\cos(2\pi nx) \cdot \cos(2\pi kx) &= \frac{1}{2}(\cos(2\pi(n-k)x) + \cos(2\pi(n+k)x)).
\end{aligned}
$$

We report in Table 10 upper and lower bounds for this ODE using trigonometric test functions. We see that the bounds are much tighter for the even moments. While the bounds are not as tight as in the earlier cases, nevertheless they do give an indication that the proposed method may have applicability beyond polynomial coefficient operators.

Table 10. Upper and lower bounds for the ODE (7) for $N = 20$ using SeDuMi

Variable	LB	UB
m_0	0.1128	3.1239
m_1	0.6730	1.0954
m_2	0.0000	0.0294
m_3	0.4471	0.7192
m_4	0.0127	0.0130
m_5	0.3349	0.5390
m_6	0.0072	0.0073
m_7	0.2678	0.4310
m_8	0.0046	0.0047
m_9	0.2231	0.3591
m_{10}	0.0032	0.0032

3.6. Insights From The Computations

In this section, we summarize the major insights from the computations we performed.

1. In both one and two dimensions, the proposed method gives strong bounds in reasonable times.
2. Perhaps the most encouraging finding is that the semidefinite constraints significantly improve over the bounds from the linear constraints. In particular, the semidefinite constraints together with only a few linear constraints, produce fairly informative bounds. These bounds are guaranteed upper and lower bounds, rather than simply approximations.
3. The semidefinite optimization solvers we used exhibited some numerical stability issues, and the semidefinite moment conditions seem to contribute to the lack of numerical stability of the black-box solvers.
4. Our experiments with trigonometric moments indicate that the proposed method is not restricted to PDEs with polynomial coefficients, but can accommodate more general coefficients by appropriately changing the underlying basis.

4. Concluding Remarks

We have presented a method for providing bounds on functionals defined on solutions of PDEs, using semidefinite optimization methods. While we present a hierarchy of increasingly tighter bounds, we emphasize that even at the initial steps, the bounds produced are, by virtue of their construction, guaranteed upper (respectively lower) bounds. Discretization based methods are unable to produce such bounds.

For the case of monomial test functions, the algorithm proposed in this paper uses degree N monomials, and hence $O(N^d)$ variables. Compared to traditional discretization methods, the proposed method provides bounds, as opposed to approximate solutions, by solving a semidefinite optimization problem on $O(N^d)$ variables. The computational results at least for one and two dimensions indicate that we obtain relatively tight bounds even with small to moderate N, which is encouraging.

Despite considerable progress in recent years, the current state of the art of semidefinite optimization codes, especially with respect to stability of the numerical calculations, is not yet at the level of linear optimization codes. This is one of the major limitations of the proposed method, as it relies on semidefinite optimization codes. Furthermore, the fact that we use polynomial moments may also be partially responsible for leading to ill-conditioned problems. It would be worth considering either different spanning families, or also different parameterizations which have better behavior. Furthermore, we use general purpose black-box semidefinite codes even though we have a very particular formulation with a lot of structure. Indeed, the moment matrices depend only on the structure of the support set, and are independent of the linear operator. The hope is that progress in the area of numerical methods for semidefinite optimization codes, and specifically codes with improved performance for moment matrices, will improve the stability and performance of the proposed method as well.

Acknowledgements. We would like to thank Professors Patera and Peraire for insightful discussions and encouragement, Professors Fujisawa, Kojima and Nakata for discussions and help regarding the SDPA code, and the associate editor and two anonymous reviewers for several insightful suggestions that improved the paper.

References

1. Akhiezer, N.: The classical moment problem, Eng. Ed., Oliver and Boyd, London 1965
2. Bertsimas, D.: The achievable region method in the optimal control of queueing systems; formulations, bounds and policies. Queueing Systems and Applications **21** (3–4), 337–389 (1995)
3. Bertsimas, D., Paschalidis, I., Tsitsiklis, J.: Optimization of multiclass queueing networks: polyhedral and nonlinear characterizations of achievable performance. Ann. Appl. Prob. **4** (2), 43–75 (1994)
4. Bertsimas, D., Popescu, I.: Optimal inequalities in probability theory: a convex optimization approach. SIAM Journal of Optimization **15**, 780–804 (2005)
5. Bertsimas, D., Popescu, I.: On the relation between option and stock prices: a convex optimization approach. Operations Research **50** (2), 358–374 (2002)
6. Bochnak, J., Coste, M., Roy, M.: Real algebraic geometry, Springer–Verlag, Berlin Heidelberg, 1998
7. Brezzi, F., Fortin, M.: Mixed and hybrid finite element methods. Springer–Verlag, New York, 1991
8. Fujisawa, K., Kojima, M., Nakata, K.: SDPA (Semidefinite Programming Algorithm) User's Manual, Version 4.10, Research Report on Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1998
9. Haviland, E.K.: On the momentum problem for distribution functions in more than one dimension II. Amer. J. Math. **58**, 164–168 (1936)
10. Harrison, J.M.: The diffusion approximation for tandem queues in heavy traffic. Adv. Appl. Prob. **10**, 886–905 (1978)
11. Harrison, J.M., Landau, H.J., Shepp, L.A.: The stationary distribution of reflected Brownian motion in a planar region. Ann. Prob. **13**, 744–757 (1985)
12. Kumar, S., Kumar, P.R.: Performance bounds for queueing networks and scheduling policies. IEEE Trans. on Aut. Control **39**, 1600–1611 (1994)
13. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**, 796–817 (2001)
14. Lasserre, J.B.: Bounds on measures satisfying moment conditions. Annals Appl. Prob. **12**, 1114–1137 (2002)
15. Paraschivoiu, M., Patera, A.: A Hierarchical Duality Approach to Bounds for the Outputs of Partial Differential Equations. Comput. Methods Appl. Eng. **158**, 389–407 (1998)
16. Paraschivoiu, M., Peraire, J., Patera, A.: A Posteriori Finite Element Bounds for Linear–Functional Outputs of Elliptic Partial Differential Equations. Comput. Meth. Appl. Mech. Eng. **150**, 289–312 (1997)
17. Parrilo, P.: Semidefinite programming relaxations for semi-algebraic problems. Math. Programming Ser. B **96**, 293–320 (2003)

18. Peraire, J., Patera, A.: Bounds for linear-functional outputs of coercive partial differential equations: Local indicators and adaptive refinement. New Advances in Adaptive Computational Methods in Mechanics, available at http://raphael.mit.edu/peraire/, 1997
19. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Differential Equations, Springer–Verlag, Berlin, 1997
20. Reznick, B.: Some concrete aspects of Hilbert's 17th problem. In: Delzell, C.N., Madden, J.J. (eds.), Real Algebraic Geometry and Ordered Structures. Cont. Math. 2000, p 253
21. Schwerer, E.: A Linear Programming Approach to the Steady–State Analysis of Markov Processes. Palo Alto, CA: Ph.D. Thesis, Stanford University, 1996
22. Schmüdgen, K.: The K–moment problem for compact semi–algebraic sets. Math. Ann. **289**, 203–206 (1991)
23. Strang, G., Fix, G.: An Analysis of the Finite Element Method, Prentice–Hall, Englewood Cliffs, NJ, 1973
24. Sturm, J.: SeDuMi Semidefinite Software, http://www/unimaas.nl/sturm
25. Trefethen, L.N., Williams, R.J.: Conformal mapping solution of Laplace's equation on a polygon with oblique derivative boundary conditions." J. Comp. and App. Math. **14**, 227–249 (1986)
26. Vandenberghe, L., Boyd, S.: Semidefinite Programming. SIAM Review **38**, 49–95 (1996)
27. Varadhan, S.R.S., Williams, R.J.: Brownian motion in a wedge with oblique reflection. Comm. Pure. Appl. Math. **38**, 405–443 (1985)
28. Watson, G.N.: A Treatise on the Theory of Bessel Functions, 2nd Edition, The Macmillan Company, Cambridge, 1994