Moment and SDP relaxation techniques for smooth approximations of problems involving nonlinear differential equations

Martin Mevissen * Jean B. Lasserre ** Didier Henrion ***

Abstract: Combining recent moment and sparse semidefinite programming (SDP) relaxation techniques, we propose an approach to find smooth approximations for solutions of problems involving nonlinear differential equations. Given a system of nonlinear differential equations, we apply a technique based on finite differences and sparse SDP relaxations for polynomial optimization problems (POP) to obtain a discrete approximation of its solution. In a second step we apply maximum entropy estimation (using moments of a Borel measure associated with the discrete solution) to obtain a smooth closed-form approximation. The approach is illustrated on a variety of linear and nonlinear ordinary differential equations (ODE), partial differential equations (PDE) and optimal control problems (OCP), and preliminary numerical results are reported.

Keywords: semidefinite programming, nonlinear optimal control, maximum entropy estimation, approximative methods, nonlinear optimization, moment methods.

1. INTRODUCTION

Problems involving nonlinear differential equations arise in a variety of models for real world problems. Even finding approximate solutions for nonlinear differential equations remains a challenge. In the previous work Mevissen et al. (2008, 2009) we have established a technique based on sparse semidefinite programming (SDP) relaxations to construct discrete approximations for solutions of systems of nonlinear differential equations. In this paper we present a novel approach to obtain smooth approximations for solutions of differential equations and of problems involving differential equations such as optimal control problems. Namely, an approximate solution is obtained by applying the maximum entropy estimation of Borwein et al. (1991); Lasserre (2007) with a finite number of moments of a Borel measure associated with a discrete approximation of the solution of the differential equation. For linear differential equations, an SDP relaxation based method was proposed in Bertsimas et al. (2006) to generate contracting sequences of lower and upper bounds for the moments.

Our contribution in this paper is primarily concerned with nonlinear differential equations and nonlinear optimal control problems. In a first step, we take advantage of the discrete approximations provided by the SDP relaxation method in Mevissen et al. (2008) to compute a finite set of moments for an appropriately defined measure with discrete support. Next, an approximation for the solution of the differential equation is obtained in closed form by maximum entropy estimation, using the moments of the discrete measure. To the best of our knowledge, it seems to be the first attempt to apply maximum entropy estimation to obtain smooth approximations for solutions of linear and nonlinear differential equations. Finally, if maximum entropy estimation does guarantee some weak convergence of the estimate to the true solution as the number of moments increases, it does not guarantee pointwise convergence on the entire domain of the differential equation. However, as our preliminary results show for different linear and nonlinear differential equations and optimal control problems, accurate pointwise approximations can be achieved on certain regions of the domain.

The structure of the paper is as follows. In Section 2 we briefly recall the sparse SDP relaxations method for solving nonlinear differential equations numerically, which is the basis for our technique. In Section 3 we introduce the

* The research of M. Mevissen was supported by the Doctoral Scholarship of the German Academic Exchange Service. The research of D. Henrion was partly supported by project No. 103/10/0628 of the Grant Agency of the Czech Republic.
method of maximum entropy estimation. Our technique to compute smooth approximations for the solutions of ordinary differential equations (ODE), partial differential equations (PDE) and optimal control problems (OCP), which combines the methods from Section 2 and 3, is presented in Section 4. Finally, preliminary numerical results for this method are reported in Section 5.

2. SPARSE SDP RELAXATIONS FOR SOLVING DIFFERENTIAL EQUATIONS

In this section we recall the approach to compute discrete approximations for solutions to systems of differential equations with polynomial data presented in Mevissen et al. (2008).

2.1 Transforming a differential equation into a POP

In this paper we consider bidimensional differential equation problems of the following type

\[ L(u(x,y)) + G(u(x,y)) = f(x,y) \forall (x,y) \in \Omega, \]

\[ H(u(x,y)) = g(x,y) \forall (x,y) \in \partial \Omega, \]

\[ \text{ldb} \leq u(x,y) \leq \text{ubd} \forall (x,y) \in \Omega, \]

with \( \Omega := [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \subset \mathbb{R}^2 \), \( f : \Omega \to \mathbb{R}^m \), \( g : \partial \Omega \to \mathbb{R}^m \), smooth functions, \( L : \mathbb{D} \to \mathbb{D} \) and \( H : \mathbb{D} \to \mathbb{D} \) differential operators, \( \mathbb{D} \subseteq L_1(\Omega) \), with \( L_1(\Omega) \) the Banach space of integrable functions on \( \Omega \), and \( \text{ldb} < \text{ubd} \) bounds for the function values of \( u \). Even if lower and upper bounds are not given by the problem directly, we need to impose them, since they are crucial for numerical stability of our method. If little is known about the solutions of (1), loose lower and upper bounds are to be imposed.

In a first step of the sparse semidefinite relaxation (SDP) method for solving differential equations, we transform an ODE or a PDE of form (1) into a polynomial optimization problem (POP). We then discretize the rectangular domain \( \Omega \) by \( N \) (one-dimensional case) or \( N_x N_y \) (two-dimensional case) grid points, approximate the derivatives by standard finite differences, for instance

\[
\frac{\partial^2 u(x,y)}{\partial x^2} \approx \frac{u(x_{i+1},y_j) - 2u(x_i,y_j) + u(x_{i-1},y_j)}{\Delta x^2},
\]

where \( u_{i,j} := u(x_i,y_j) \), and denote the discretized differential operators at some interior grid point \( (x_{i},y_j) \) and boundary grid point \( (x_{k},y_l) \) as \( L_{i,j} \) and \( H_{k,l} \), respectively. Finally, we take the discretization of the problem (1) as a system of constraints, choose some objective function \( F \) which needs to be polynomial in \( u \), i.e. \( F(u) \in \mathbb{R}[u] \), and obtain the following optimization problem

\[
\min F(u) \quad \text{s.t.} \quad L_{i,j}(u) + G(u_{i,j}) = f(x_i,y_j), \quad \text{ldb}_{i,j} \leq u_{i,j} \leq \text{ubd}_{i,j} \forall (i,j) \in \{1, \ldots, N_x\} \times \{1, \ldots, N_y\},
\]

\[
H_{k,l}(u) = g(x_k,y_l) \forall (k,l) \in \{1, \ldots, N_x\} \times \{1, \ldots, N_y\} \cup \{1, \ldots, N_x\} \times \{1, \ldots, N_y\}
\]

which, since all functions are polynomial in the variable \( u = (u_{1,1}, \ldots, u_{N_x,N_y}) \), is a POP. In the case of an OCP the objective function \( F \) is given by the discretization of the optimal value function. In the case of a differential equation with several solutions we can pick a specific solution by choosing an appropriate objective function. A particular choice for the objective function may be motivated by the underlying problem, such as energy functionals in physics.

2.2 Sparse SDP relaxations

Problem (2) is of the form:

\[ \text{POP} \min F(u) \quad \text{s.t.} \quad g_j(u) \geq 0 \quad \forall j \in \{1, \ldots, k\}, \]

\[ h_i(u) = 0 \quad \forall i \in \{1, \ldots, l\}, \]

\[ \text{ldb}_s \leq u_s \leq \text{ubd}_s \forall s \in \{1, \ldots, n\}, \]

where \( F(u) = \sum_{\alpha \in \mathbb{N}^n} \alpha ! F_{\alpha} u^\alpha \), \( g_j(u) = \sum_{\alpha \in \mathbb{N}^n} g_{j,\alpha} u^\alpha \) and \( h_i(u) = \sum_{\alpha \in \mathbb{N}^n} h_{i,\alpha} u^\alpha \). As mentioned above, the dimension of this POP is \( \bar{n} = N_x N_y \) (or \( \bar{n} = N \) in the one-dimensional case). But due to the structure of the finite difference discretization only a small number of the \( \bar{n} \) components of \( u \) occur in each constraint or in each monomial of the objective function of (2). Thus, the POPs derived from nonlinear differential equations are sparse.

A systematic way of characterizing the structured sparsity of a POP has been introduced in Waki et al. (2006), and this structured sparsity is exploited by the sparse SDP relaxations constructed in Waki et al. (2006). For the POP (3) one obtains the sparse SDP relaxations:

\[ \text{SDP}_w \min \sum_{|\alpha| \leq w} F_{\alpha} y_{\alpha} \quad \text{s.t.} \quad M_{w-w}(y_{\alpha}, (1) \leq \bar{n}) \geq 0 \quad \forall j \in \{1, \ldots, k\}, \]

\[ M_{w-w}(y_{\alpha}, (1) \leq \bar{n}) = 0 \quad \forall i \in \{1, \ldots, l\}, \]

\[ \text{ldb}_s \leq y_s \leq \text{ubd}_s \forall s \in \{1, \ldots, n\}, \]

where \( w := \lceil \deg g_j \rceil, \quad w' := \lceil \deg h_i \rceil, \quad (1) \leq \bar{n} \) is the set of subsets of \( \{1, \ldots, \bar{n}\} \) derived from the correlative sparsity pattern matrix of the POP Waki et al. (2006), \( M_{w-w-}(-, -) \) are the partial moments and localizing matrices of Lasserre (2006), and \( w \geq w \min := \max \{1, \ldots, \bar{n}\} \) is the order of the SDP relaxation. As \( w \) increases, solving SDP \( w \) generates a nondecreasing sequence of lower bounds for min(POP), namely:

\[ \min (\text{SDP}_w) \leq \min (\text{SDP}_{w+1}) \leq \cdots \leq \min (\text{POP}). \]

Moreover, if (3) has a unique minimizer \( x^* \), the vector \( (y_{\alpha,1}^*, \ldots, y_{\alpha,n}^*) \) obtained from an optimal solution \( y_{\alpha}^* \) of (4) is an approximation of \( x^* \), and under certain compactness conditions, see Lasserre (2006), as \( w \to \infty \),

\[ \min (\text{SDP}_w) \to \min (\text{POP}) \quad \text{and} \quad y_{\alpha}^* \to x^*. \]

In implementations of the sparse SDP relaxation such as SparsePOP, small linear perturbation terms are added to the objective function \( F \), in order to ensure that (2) has a unique optimal solution. Therefore, an optimal solution of the SDP (4) is an approximation of the optimal solution of the POP (2).

2.3 Discrete approximation

From (5), asymptotic convergence of minimum and minimizer of SDP \( w \) to minimum and minimizer of POP (3) are guaranteed under uniqueness of the optimal solution and compactness of the feasible set of the POP. However, recall that the dimension \( \bar{n} \) of POP (3) is \( N_x N_y \), and so, if \( w \) is much larger than \( w \ min \), the resulting SDP \( w \) in (4)
is untractable for general purpose SDP solvers. Therefore, we have to restrict ourselves to \( u \in \{ w^\text{min}, w^\text{min} + 1 \} \) for medium scale \( N \), and we cannot expect that SDP \( w^\text{min} \) or SDP \( w^\text{max} + 1 \) provide accurate approximations to a solution of a nonlinear differential equation. In order to improve accuracy, we may apply additional locally convergent optimization techniques. For instance sequential quadratic programming (SQP) can be applied to (2), starting from the solution of SDP \( w^\text{0} \), as initial guess. Since the solution of the SDP relaxation is an approximation for the global optimizer, it is a systematic choice for an initial point for local methods which does not require any a priori information. Combining the sparse SDP relaxation with SQP (or another local method) is summarized in the scheme:

**Method 1. The SDPR method**

1. Choose a discretization \((N_x, N_y)\) for the differential equation problem.
2. Choose a relaxation order \( w \) and an objective function \( F \). Apply SDP \( w \) to (2) and obtain its solution \( u^\star \) and the lower bound solution \( u \) and the lower bound solution \( u \).
3. Apply sequential quadratic programming (SQP) to (2) with \( u^\star \) as initial guess, and obtain \( u \) as discrete approximation to a solution of the differential equation problem.

### 3. MAXIMUM ENTROPY ESTIMATION

In this section we briefly introduce the maximum entropy estimation of Borwein et al. (1991); Lasserre (2007, 2009), our second tool to find smooth approximations for solutions of nonlinear differential equations. The maximum entropy estimation is concerned with the following problem: Let \( u \in L_1(\Omega) \) be nonnegative and partially known by the finite vector \( m \) of moments up to order \( M \) of the associated Borel measure \( d\mu := u \, dx \, dy \) on \( \Omega \). From the only knowledge of \( m \), find an estimate \( u^\star_M \in L_1(\Omega) \) such that all moments of order up to \( M \) of the measure \( d\mu_M := u^\star_M \, dx \, dy \) match those of \( d\mu \) and analyze the asymptotic behavior of \( u^\star_M \) as \( M \to \infty \). An elegant method consists in finding the estimate \( u^\star_M \) that maximizes the Boltzmann-Shannon entropy functional

\[
 u \mapsto \int_\Omega u \ln u \, dx \, dy, \quad u \in L_1(\Omega).
\]

In this case, the optimal estimate \( u^\star_M \) is given by

\[
 u^\star_M(x, y) = \tilde{u}^\star_M(v^\star, x, y) = \exp \left( \sum_{0 \leq i,j \leq M} v^\star_i x^i y^j \right),
\]

where \( v^\star \in \mathbb{R}^{[m]} \) is a (global) optimizer of the convex finite-dimensional optimization problem:

\[
 \max_{v \in \mathbb{R}^{[m]}} \{ m, v \} - \int_\Omega u^\star_M(v, x, y) \, dx \, dy.
\]

**Proposition 1.** Let \( u^\star_M \) be obtained from an optimal solution of (7). Then, as \( M \to \infty \),

\[
 \int_\Omega \psi(x, y) u^\star_M(x, y) \, dx \, dy \rightarrow \int_\Omega \psi(x, y) u(x, y) \, dx \, dy,
\]

for every bounded measurable function \( \psi : \Omega \to \mathbb{R} \) which is continuous almost everywhere.

However, the pointwise convergence \( u^\star_M(x, y) \to u(x, y) \) does not hold in general.

### 4. SMOOTH APPROXIMATIONS TO SOLUTIONS OF DIFFERENTIAL EQUATIONS

In this section we show how to combine the SDPR method and maximum entropy estimation to obtain smooth approximations for solutions of linear and nonlinear differential equations. Our discussion focuses on the 2-dimensional case, but the 1-dimensional case is covered analogously. Let \( u \in D \) be a solution of (1) and assume without loss of generality that \( \text{lb}d \geq 0 \) so that \( u \) is nonnegative on \( \Omega \).

For \( \text{lb}d < 0 \) define \( \tilde{u} := u - \text{lb}d \) and apply the outlined procedure to the new function \( \tilde{u} \). Associated with \( u \), let \( d\mu(x, y) := u \, dx \, dy \) be the finite Borel measure on \( \Omega \) with moment vector \( m = (m_{i,j}) \) of all moments up to order \( M \):

\[
 m_{i,j} := \int_\Omega x^i y^j \, d\mu(x, y) = \int_\Omega x^i y^j u(x, y) \, dx \, dy,
\]

for \( (i, j) \in \mathbb{N}^2 \) with \( i + j \leq M \). For linear PDEs, a hierarchy of tightening lower and upper bounds for the components \( m_{i,j} \) of the moment vector can be obtained by solving a sequence of SDPs as proposed in Bertsimas et al. (2006). However, this approach cannot be applied in the case of nonlinear PDEs. Our strategy is to use a discrete approximation \((u_{k,l})_{1 \leq k \leq N_x, 1 \leq l \leq N_y}\) of a solution of the PDE to approximate the moments \( m_{i,j} \) by:

\[
 m_{i,j}^\Delta := \sum_{k=1}^{N_x} \sum_{l=1}^{N_y} x_k^i y_l^j u_{k,l} \Delta x \Delta y,
\]

for \( (i, j) \in \mathbb{N}^2 \) with \( i + j \leq M \), where \( \Delta x := \frac{2\text{lb}d - \text{ub}d}{N_x - 1} \) and \( \Delta y := \frac{2\text{ub}d - \text{lb}d}{N_y - 1} \). If the discretization \((N_x, N_y)\) is sufficiently fine and \((u_{k,l})_{k,l}\) is a close approximation of \( u \), then we expect \( m_{i,j}^\Delta \) to be a good approximation of \( m_{i,j} \) for all \( (i, j) \in \mathbb{N}^2 \). Thus, we can apply maximum entropy estimation to the vector \( m^\Delta \) to obtain a smooth approximation of \( u \). This idea is formalized in the following algorithm for obtaining smooth approximations of solutions of linear or nonlinear PDE problems.

**Method 2. The smooth SDP approximation method**

Given a PDE problem of form (1).

1. Choose a discretization \((N_x, N_y)\), relaxation order \( w \) and objective \( F \), and apply the SDPR method to obtain a discrete approximation \((u_{k,l})_{k,l}\) to a solution of the PDE problem. If not given in the formulation of the PDE problem, impose lower and upper bounds, \( \text{lb}d \) and \( \text{ub}d \) for \( u \).
2. Choose a moment bound \( M \in \mathbb{N} \) and use \((u_{k,l})_{k,l}\) to calculate \( m^\Delta \) as in (8).
3. Apply maximum entropy estimation to \( m^\Delta \) and obtain vector \( v^\star \in \mathbb{R}^{[m^\Delta]} \), optimal solution of (7).
(4) Obtain the approximation $u^*_M$ with $u^*_M(x, y) = \sum_{i,j} \exp(v^*_i y^j)$ for a solution $u$ of the PDE problem (1) on $\Omega$.

As $M \to \infty$, $u^*_M \to u$ weakly (see Proposition 1) but not pointwise, i.e., one cannot guarantee $u^*_M(x, y) \to u(x, y)$ on $\Omega$. Nevertheless, as reported in Lasserre (2007) the maximum entropy estimation may provide accurate pointwise approximation of the unknown function to be recovered on certain segments of the domain $\Omega$. We next illustrate on a variety of PDE problems and OCPs, that indeed good pointwise approximation can be obtained on some parts of the domain $\Omega$.

5. NUMERICAL EXPERIMENTS

We illustrate the potential of Method 2 on a variety of ODE and PDE problems. As an implementation of the sparse SDP relaxations we apply the software SparsePOP of Waki et al. (2005) and as an implementation of sequential quadratic programming (SQP) in the SDPR method (7) we apply the Matlab Optimization Toolbox commands fminunc and fmincon, respectively. As we restrict ourselves to ODEs and PDEs with rectangular domains $\Omega$ we can apply standard quadrature and cubature formulas to compute the gradient and Hessian for (7). Thus, we apply the Matlab commands trapz in the one-dimensional case and dblquad in the two-dimensional case, respectively. In order to evaluate the quality of the smooth approximation provided by Method 2 we define the average error $\epsilon_u(M) := \Delta x \sum_i u(x_i) - u^*_M(x_i)$ and the maximum pointwise error $\epsilon_{\max}(M) := \max_i | u(x_i) - u^*_M(x_i) |$.

5.1 Linear differential equations

As first test problems for our approach we consider a linear ODE and a linear PDE from Bertsimas et al. (2006) and compare (a) the discretized moment approximations obtained by Method 2 to the bounds obtained in Bertsimas et al. (2006), and (b) the smooth approximation $u^*_M$ to the known analytic solution.

**Linear ODE** The linear ODE is given by:

\[
u''(x) + 3u'(x) + 2u(x) = 0 \quad \forall x \in [0, 1],
\]

\[u'(0) = -2e^2, \quad u(1) = -2. \tag{9}\]

For this problem the unique solution is $u(x) = e^2 e^{-2x}$. We apply Method 1 with $F = \sum u_i$, $N = 2000$ and $w = 1$. We calculate the approximate moments $m\Delta$ for $M = 40$ and compare them to approximate moments for $m$ derived in Bertsimas et al. (2006) by contracting lower and upper bounds in Table 1. All moments coincide up to the fourth digit.

We apply Method 2 for $M \in \{1, \ldots, 5\}$ and report the resulting vector $v^*_M$ in Table 2.

Note that the actual solution of the ODE problem corresponds to a parameter vector $v_{\text{opt}} = (2, -2, 0, \ldots)$ as reported in Table 2, $(v^*_i(M), v^*_j(M)) \approx (2, -2)$, and even though $v^*_i(M) \neq 0$ for $i > 1$, the maximum pointwise error is quite small. However, $(v^*_i(M), v^*_j(M))$ does not converge to $(2, -2)$ since pointwise convergence is not guaranteed.

**Linear PDE** The linear PDE is given by

\[
u_{xx}(x, y) + 3\nu_{xy}(x, y) - 3e^{x+y} = 0 \quad \forall (x, y) \in [0, 1]^2, \tag{10}\]

where the boundary conditions are set up such that $u(x, y) = e^{x+y}$ is the unique solution of the PDE problem. We apply the SDPR method with $F(u) = -\sum_{k,l} u_{k,l}$, $w = 1$ and $N_x = N_y = 100$, compute $m\Delta$ for all moments of order up to 3 and compare these moments to the exact moments derived from the known solution and to the best lower and upper bounds for the moments from Table 7 in Bertsimas et al. (2006). See Table 3 for the results. Our approach provides some approximations for the moments that is well within the bounds LB and UB from Bertsimas et al. (2006).

In fact the higher moment terms $v^*_i(M)$ for $i > 1$ counterbalance the difference between $(v^*_0(M), v^*_1(M))$ and $(2, -2)$.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$m_{\Delta}^0$</th>
<th>$m_{\Delta}^1$</th>
<th>$m_{\Delta}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.1942</td>
<td>3.1945</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.0957</td>
<td>1.0973</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.0886</td>
<td>1.0888</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.0624</td>
<td>0.0624</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.0256</td>
<td>0.0256</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Approx. moments $m\Delta$ for $N = 2000$ compared to approx. $m_{\Delta}^{BC}$ derived from lower and upper bounds in Bertsimas et al. (2006) for linear ODE (9)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$m_{\Delta}^0$</th>
<th>$m_{\Delta}^1$</th>
<th>$m_{\Delta}^2$</th>
<th>$m_{\Delta}^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.0043</td>
<td>2.0042</td>
<td>2.0043</td>
<td>2.0039</td>
</tr>
<tr>
<td>1</td>
<td>-2.0135</td>
<td>-2.0083</td>
<td>-1.9922</td>
<td>-1.9744</td>
</tr>
<tr>
<td>3</td>
<td>0.0851</td>
<td>0.2886</td>
<td>0.3629</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.1180</td>
<td>0.2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-0.2604</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Smooth SDP approx. for linear ODE (9)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$m_{\Delta}^0$</th>
<th>$m_{\Delta}^1$</th>
<th>$m_{\Delta}^2$</th>
<th>$m_{\Delta}^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0042</td>
<td>0.0044</td>
<td>0.0042</td>
<td>0.0039</td>
</tr>
<tr>
<td>1</td>
<td>0.0242</td>
<td>0.0228</td>
<td>0.0226</td>
<td>0.0222</td>
</tr>
<tr>
<td>3</td>
<td>0.0242</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Approx. moments for $N_x = N_y = 100$ compared to exact moments and bounds from Bertsimas et al. (2006) for linear PDE (10)
For both linear differential equations we obtain accurate approximations of the moments for the measures associated with the unique solutions. Moreover, unlike Bertsimas et al. (2006), we can exploit these moment approximation to find smooth approximations for the actual solutions.

5.2 A nonlinear elliptic PDE

As a first example of an optimal control problem is given by the elliptic PDE:

\[
\begin{align*}
& u_{xx} + u_{yy} + 22u(1 - u^2) = 0 \quad \text{on } [0, 1]^2, \\
& u = 0 \quad \text{on } \partial(0, 1)^2, \\
& 0 \leq u \leq 1 \quad \text{on } [0, 1]^2.
\end{align*}
\]

This problem is well known to have a nontrivial positive solution. We apply Method 2 with \( F(u) := -\sum_{k,l} u_{x,l} \), \( w = 2 \), \( N_x = N_y = 49 \) and \( M \in \{2, 3\} \), and obtain an accurate discrete approximation \((u_{i,j})_{i,j}\) and two smooth approximations \(u_2^\star\) and \(u_3^\star\) with errors \((\epsilon_{u_{\max}}(2), \epsilon_u(2)) = (0.081, -0.002)\) and \((\epsilon_{u_{\max}}(3), \epsilon_u(3)) = (0.076, -0.002)\), respectively. In Figure 1 it is illustrated how accurate the smooth approximation \(u_3^\star\) resembles the shape of the discrete approximation \((u_{i,j})_{i,j}\).

5.3 Reaction diffusion equation

A challenging ODE problem which is known to have many solutions is given in Mimura (1979) by:

\[
\begin{align*}
& -\frac{1}{20} u_{xx} + \frac{1}{9} (35 + 16u - u^2) \\ & \quad u_{xx} + 22u(1 - u^2) \\ & \quad 0 \quad \text{on } [0, 5], \\
& 4v_{xx} - \left(1 + \frac{2}{5} v\right) u_{xx} + 2uv = 0 \quad \text{on } [0, 5], \\
& u'(0) = u'(5) = v'(0) = v'(5) = 0, \\
& 0 \leq u, v \leq 14 \quad \text{on } [0, 5].
\end{align*}
\]

To ensure numerical stability we scale the domain \([0, 5]\) of (12) to \([0, 1]\) before applying Method 2, as \(\exp(x^2)\) gets very large for \(|x| > 1\). Problem (12) involves two functions \(u\) and \(v\), thus we need to apply Method 2 twice, once for \(u\) and once for \(v\), in order to obtain smooth approximations for both functions. Note that the SDPR method only needs to be applied once and provides a discrete approximation \((u_{i,j})_{i,j}\), maximum entropy estimation needs to be applied twice. We apply Method 2 for \(F(u, v) = -u_{\frac{x}{2}}\), \(N = 100\), \(w = 3\) and \(M \in \{10, 20, 30, 50\}\). Solutions for \(u\) and \(v\) and the derived smooth approximations are pictured in Figure 2 and the errors are reported in Table 5. Even though the average and the maximum pointwise approximation errors are not small, even for large \(M\), we observe that the pointwise approximation gets more and more accurate on certain segments of the domain - in this case the third part of the interior, where the third peak of \(u\) and \(v\) occurs. The pointwise approximation is not good - or does even get worse - near the boundary of \([0, 1]\).

5.4 Control of production and consumption

A first example of an optimal control problem is given by

\[
\begin{align*}
& \min - \int_0^T (1 - u(t))x(t) \, dt \\
& \text{s.t.} \quad x(t) = u(t)x(t) \quad \forall t \in [0, T], \\
& \quad x(0) = x_0, \\
& \quad 0 \leq x(t) \leq 1 \quad \forall t \in [0, 1], \\
& \quad 0 \leq x(t) \leq 10 \quad \forall t \in [0, 10],
\end{align*}
\]

where \(T > 1\) fixed. For this simple problem the optimal control law is given by

\[
u^*(t) = \begin{cases} 
1 & \text{if } 0 \leq t \leq T - 1, \\
0 & \text{if } T - 1 < t \leq T.
\end{cases}
\]

We choose \(x_0 = 0.25, T = 4\) and \(N = 100\), and apply Method 2 after scaling the domain to \([0, 1]\) for \(M \in \{5, 10, 20, 30\}\). In this case the objective function \(F\) is given by a discretization of the objective function of (13). As in the case of ordinary and partial differential equations, we observe that a fairly good pointwise approximations of both, optimal control law and corresponding trajectory, are obtained on the interior of the domain. See Figure 3.

5.5 The double integrator

Another interesting control problem, which has been discussed in Lasserre et al. (2008), is given by

\[
\begin{align*}
& \min T \\
& \text{s.t.} \quad \dot{x}_1(t) = x_2(t) \quad \forall t \in [0, T], \\
& \quad \dot{x}_2(t) = u(t) \quad \forall t \in [0, T], \\
& \quad x(0) = x_0 \in \mathbb{R}^2, \\
& \quad x(T) = (0, 0), \\
& \quad -1 \leq u(t) \leq 1 \quad \forall t \in [0, T], \\
& \quad -1 \leq x_1(t), x_2(t) \leq 10 \forall t \in [0, T].
\end{align*}
\]

After scaling the domain to \([0, 1]\), we apply Method 2 with \(N = 50, \omega = 3, x_0 = (0.8, -1)\) and \(M \in \{5, 10, 20, 30\}\). Since \(T\) is not fixed, it is treated as an additional variable, i.e., we consider the polynomial optimization problem with variable \((x_1, x_2, x_3, x_4, x_5, u_1, \ldots, u_N, T)\). We observe that the pointwise approximation of optimal control and trajectories on the interior of the domain gets better and better for increasing moment order, see Figure 4. Note, unlike the moment based method in Lasserre et al. (2008) which yields bounds for the optimal value of the control problem, we obtain discrete and smooth approximations to the optimal control and the corresponding trajectories by Method 1 and Method 2, respectively.

6. CONCLUDING REMARKS

We introduced a novel technique to derive smooth approximations for solutions of systems of differential equations and optimal control problems, which is based on sparse SDP relaxations and the maximum entropy estimation method. As demonstrated on some examples of nonlinear ordinary differential equations, partial differential equations and optimal control problems, this technique is
Fig. 1. Discrete approx. (left), smooth approx. (center) for the same grid, and pointwise difference between the two (right, notice the vertical axis scale) for nonlinear PDE (11).

Fig. 2. Smooth approx. for $u$ (left) and $v$ (right) for $M \in \{10, 20, 30, 50\}$ for nonlinear ODE (12).

Fig. 3. Smooth approx. for $x$ (left) and $u$ (right) for $M \in \{5, 10, 20, 30\}$ for OCP (13).

Fig. 4. Smooth approx. for $x_1$ (left), $x_2$ (center) and $u$ (right) for $M \in \{5, 10, 20, 30\}$ for OCP (14).
promising to obtain accurate pointwise approximations of a solution of a differential equation on segments of its domain. It would be interesting to characterize regions of the domain of nonlinear differential equations where accurate pointwise approximations can be guaranteed. Another question is whether a different, less restrictive parametrization than (6) allows better pointwise approximations on larger segments of the domain. Of course, other choices of entropy would lead to different results, and it remains a topic of future investigation to analyze, whether or not some choice should be preferred. At this point solving the sparse SDP relaxation is the major computational bottleneck in Method 2. Every improvement for approximation accuracy and efficiency of SDP relaxation techniques for polynomial programs and every improvement for SDP solvers will extend the range of the proposed method.

In order to validate these techniques in the present context of PDEs, a comparison with state-of-the-art numerical methods for solving nonlinear PDEs remains to be done. In particular, the maximum entropy technique should be compared with standard and spline interpolation methods which both provide differentiable approximations that coincide with the discrete solution at each grid point. But, unlike our method, in polynomial interpolation the degree of the polynomial is directly related to the number of grid points. Also, cubic splines are twice differentiable at the grid points, whereas our approximation is smooth. In contrast, and instead of searching for a smooth solution that matches the unknown \( u \) at the grid points, one searches for a smooth solution that matches finitely many moments of the associated measure \( d\mu = udx \). And so, an interesting feature is to obtain a smooth approximation from a limited number of moments and not from the discrete solution explicitly (still, before using moments, a sufficiently fine discretization is necessary to ensure \( m^\Delta \approx m \)). This method is an attempt to deal with the curse of dimensionality in the numerical analysis of PDEs. Furthermore, we emphasize that the proposed method can be applied to a wide range of problems that encompasses nonlinear partial differential equations and nonlinear optimal control problems. A detailed comparison is however out of the scope of this paper, whose main objective was to pave the way for the development of SDP techniques to solve optimal control design problems for dynamical systems described by polynomial ODEs and PDEs, in the spirit of Lasserre et al. (2008).

REFERENCES


