An Alternative Approach for Non-Linear Optimal Control Problems Based on the Method of Moments

René Meziat, Diego Patiño

Departamento de Matemáticas, Universidad de los Andes, Carrera 1 No. 18A-70, Bogotá, Colombia rmeziat@uniandes.edu.co Pablo Pedregal

E.T.S. de Ingenieros Industriales, Universidad de Castilla La Mancha Edificio Politécnico, Ciudad Real, 13071, Spain.

Abstract

We propose an alternative method for computing effectively the solution of non-linear, fixed-terminal-time, optimal control problems when they are given in Lagrange, Bolza or Mayer forms. This method works well when the nonlinearities in the control variable can be expressed as polynomials. The essential of this proposal is the transformation of a nonlinear, non-convex optimal control problem into an equivalent optimal control problem with linear and convex structure. The method is based on global optimization of polynomials by the method of moments. With this method we can determine either the existence or lacking of minimizers. In addition, we can calculate generalized solutions when the original problem lacks of minimizers. We also present the numerical schemes to solve several examples arising in science and technology.

1 Introduction

In this work we are concerned with non-linear, optimal control problems with fixed terminal time, given in the general Lagrange or Bolza forms:

$$\min_{u} \int_{0}^{1} f(x, t, u) dt + F(x(1))$$
s.t. $\dot{x} = g(x, t, u)$

$$x(0) = x_{0}$$

$$x \in \mathbb{R}^{n}, \quad u \in \mathbb{R},$$
(1)

where the functions f and g can be expressed as polynomials in the control variable. In general we have:

$$f(x,t,u) = \sum_{k=0}^{N_1} a_k(x,t)u^k$$

$$g(x,t,u) = \sum_{k=0}^{N_2} c_k(x,t)u^k$$
(2)

where the highest degree $K = \max\{N_1, N_2\}$ must be even and $a_{N_1}, c_{N_2} > 0$ in order to guarantee the coercivity of the Hamiltonian of the problem (1). We assume existence and continuity of f_x , g_x , F_x and continuity on f, g and F. Although expression (1) refers to optimal control problems in Lagrange and Bolza forms, all the arguments and results of our proposal also work well on optimal control problems given in Mayer form. The non-linear, non-convex form of the control variable, prevents us to use the Hamilton equations of the maximum principle and non-linear mathematical programming techniques on them. Both approaches would entail severe difficulties, either in the integration of the Hamilton equations or in the search method of any numerical optimization algorithm. As an alternative approach for dealing with this kind of problems we propose to convexify the control variable by using the method of moments in the polynomial expressions (2). Thus, we introduce a linear, convex relaxation of (1) in the following form:

$$\min_{m(t)} \int_{0}^{1} \sum_{k=0}^{N_{1}} a_{k}(x,t) m_{k}(t) dt + F(x(1))$$

s.t. $\dot{x} = \sum_{k=0}^{N_{2}} c_{k}(x,t) m_{k}(t)$
 $x(0) = x_{0}$ (3)

where the new control variable is the vector m, whose entries must represent the algebraic moments of a probability measure supported on the real line of the control variable u. See [1, 2, 16, 29, 30, 31, 53] for classical and modern results on the problem of moments and [17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 53, 54, 55, 56, 57, 58, 59, 60, 61] for recent results on global optimization of polynomial expressions by the characterizations of moments. Convexification of non-convex, optimal control problems by using probability measures is a well known tool exploited by other authors [3, 4, 5, 8, 9, 10, 11, 12, 13, 14]. What is new in this approach is the convexification of the control variable by using moment variables, which allow us to obtain an equivalent, convex formulation more appropriated to be solved by high performance numerical computing. It is striking the fact that the solution of the convex, discrete, formulation provides important information about the existence and behavior of the minimizers of the original non-linear, optimal control problem. Indeed, as formulation (3) takes the form of a convex, optimal control problem, we can certify the existence of minimizers and estimate them numerically by mathematical programming. In this work we will characterize the existence of minimizers of (1) according to one particular algebraic form in the minimizers of (3). In this way we obtain a practical method for estimating minimizers of (1) and at the same time, a practical method for certifying either existence or lacking of minimizers of (1)-(2). Finally, we stress the importance of the computational aspects of our proposal for high performance optimization. On one hand, we must use very specialized computational tools for obtaining numerical solutions of the relaxed model (3). We address this point in Section 4. On the other hand, by following this proposal, practitioners will be able to model and simulate many non-linear situations in science and engineering. We present a few examples of them in Section 4. Thus, we pave the way for analise, solve and simulate important families of non-linear, optimal control problems like non-holonomic constraints in robotics, sub-Riemannian geodesics and Lie brackets in geometric control, sliding modes and solid phase microstructures in structural design. See [46, 47, 48, 49, 50, 51, 52, 62].

The present paper is organized as follows. In Section 2 we describe the convexification of polynomial expressions by the method of moments and its implications in the analysis of optimal control problems given in the form (1) - (2). Section 3 explains the transformation of these problems into the equivalent formulation (3). We give necessary and sufficient conditions for the existence of minimizers of (1) - (2) by using particular features of the minimizers of the convex formulation (3). In Section 4 we solve several examples which entail high demanding numerical tasks. We finish with a short conclusion in Section 5. Section 6 includes a short afterword about the new born community on optimization of polynomials with moments.

2 Convexification of polynomial expressions

Let f and g be the polynomials shown in (2), so the Hamiltonian H of the optimal control problem (1) must have a polynomial form in the control variable u:

$$H = H(x, t, p, u) = \sum_{k=0}^{K} \alpha_k(x, t, p) u^k$$
(4)

where $K = \max\{N_1, N_2\}$. Thus, the global minimization of H in u:

$$\min_{u} H\left(u\right) = \sum_{k=0}^{K} \alpha_{k} u^{k} \tag{5}$$

is a problem well suited to be solved by the method of moments [17, 18, 19, 20, 21, 22, 24, 25, 54, 56, 57, 58, 59, 60, 61]. The essentials of this method follow.

For solving non-convex polynomial programs like (5), we can use the convex hull of the graph of the polynomial H provided it be a coercive function, that is: $\alpha_K > 0$ with even K. We can describe such convex set in the following way:

$$\operatorname{co}\left(\operatorname{graph}(H)\right) = \left\{\int_{R} \left(u, H\left(u\right)\right) \, d\mu\left(u\right) \, : \, \mu \in P\left(R\right)\right\},\tag{6}$$

where P(R) stands for the family of all probability Borel measures supported in the real line.

Theorem 1 Let H(u) be an even degree, algebraic polynomial whose leader coefficient α_K is positive, then we can express the convex hull of the graph of H as given in (6).

To prove this result, apply the separation theorem of convex analysis. Once we have characterised the convex hull of the graph of H, we can obtain the set of all global minima of H by noticing that:

$$\operatorname{argmin}(H) \subseteq \operatorname{argmin}(H_c)$$

where H_c stands for the convex envelope of H. Since H is a coercive polynomial, notice that

$$\operatorname{co}(\operatorname{graph}(H)) = \operatorname{Epigraph}(H_c)$$

Then, we can pose the global optimization problem (5) as the following optimization problem defined in probability measures:

$$\min_{u \in P(R)} \int_{R} H(u) \ d\mu(u) \tag{7}$$

whose solution is the family of all probability measures supported in argmin(H). See [17]. **Theorem 2** [17, 24, 25] When H is coercive, the set of solutions of (7) is the set of all probability measures supported in the set of global minima of H, i.e. argmin(H).

Corollary 3 [17, 24, 25] When argmin(H) is the singleton $\{u^*\}$, the Dirac measure: $\mu^* = \delta_{u^*}$ is the unique solution of (7).

Now we use the polynomial structure of the objective function H in order to transform the optimization problem (7) into the following optimization problem:

$$\min_{m \in M} \sum_{k=0}^{K} \alpha_k m_k$$

where M is the convex set of all vectors in \mathbb{R}^{K+1} whose entries are the algebraic moments of a probability measure supported in the real line. Although this formulation seems very attractive due to the linear form of the objective function and the convex structure of the feasible set, it is still a theoretical formulation not very useful if we do not properly characterize the feasible set composed of moment vectors: M. However, this is precisely the question of the classical Truncated Hamburger Moment Problem [1, 2, 16]. Its solution is easily summarized as follows: the closure of M is composed of all vectors in \mathbb{R}^{K+1} whose entries form a positive semidefinite Hankel matrix [16, 23, 27, 29, 30]:

$$\overline{M} = \{ (m_i)_{i=0}^K \in \mathbb{R}^{K+1} : (m_{i+j})_{i,j=0}^{\frac{K}{2}} \quad \text{is positive semidefinite with } m_0 = 1 \}.$$

This result allows us to transform the problem (7) into the mathematical program:

$$\min_{m} \sum_{k=0}^{K} \alpha_k m_k \quad \text{s.t.} \quad (m_{i+j})_{i,j=0}^{\frac{K}{2}} \ge 0, \text{ with } m_0 = 1$$
(8)

which has the form of a semidefinite program. See [23, 38] for an introduction to conic and semidefinite programming.

Theorem 4 [24, 25] When H is a coercive polynomial, the set of solutions of (8) is the set of all vectors $m^* \in \mathbb{R}^{K+1}$ whose entries are the algebraic moments of some probability measure supported in $\operatorname{argmin}(H)$, which is a finite set with $\frac{K}{2}$ points at the most.

Corollary 5 [17] When H is a coercive polynomial with a unique global mimimum u^* , the program (8) has a unique solution $m^* \in \mathbb{R}^{K+1}$ composed by the algebraic moments of the Dirac measure δ_{u^*} . Thus, $m_k^* = (u^*)^k$, for $k = 0, \ldots, K$.

Hence, the global minimization of the Hamiltonian H can be formulated as:

$$\min_{m} \widetilde{H}(x,t,p,m) = \sum_{k=0}^{K} \alpha_k (x,t,p) m_k \quad \text{s.t.} \quad (m_{i+j})_{i,j=0}^{\frac{K}{2}} \ge 0, \text{ with } m_0 = 1$$
(9)

where the variables x, t and p are fixed. Notice that any solution $m^*(x, t, p)$ of (9) is composed of the algebraic moments of some probability measure supported in argmin(H(x, t, p)). Since argmin(H(x, t, p)) is finite, $m^*(x, t, p)$ can be expressed as

$$m^* = \sum_{i=1}^{K'} \lambda_i (1, v_i, v_i^2, \dots, v_i^K)$$
(10)

where $argmin(H(x,t,p)) = \{v_1(x,t,p), v_2(x,t,p), \ldots, v_{K'}(x,t,p)\}$ with $K' \leq \frac{K}{2}$. Therefore, if argmin(H) is the singleton $\{u^*(x,t,p)\}$, the optimal control can be expressed as:

$$u^{*}(x,t,p) = m_{1}^{*}(x,t,p)$$

because the entries of $m^*(x, t, p)$ are the moments of the Dirac measure $\delta_{u^*(x,t,p)}$. In this work we will solve the non-linear, non-convex problem (1)-(2) by working out its convex relaxation:

$$\min_{m(t)} \int_{0}^{1} \sum_{k=0}^{N_{1}} a_{k}(x, t) m_{k}(t) dt + F(x(1))$$
s.t. $\dot{x} = \sum_{k=0}^{N_{2}} c_{k}(x, t) m_{k}(t)$

$$x(0) = x_{0}$$

$$(m_{i+j}(t))_{i,j=0}^{\frac{K}{2}} \ge 0, \text{ with } m_{0}(t) = 1$$

$$\forall t \in (0, 1).$$
(11)

We highlight the fact that the semidefinite program (9) corresponds to the optimization of the Hamiltonian of the convex formulation (11):

$$\widetilde{H} = \widetilde{H}(x, t, p, m) = \sum_{k=0}^{N_1} a_k(x, t) m_k + p^t \cdot \sum_{k=0}^{N_2} c_k(x, t) m_k = \sum_{k=0}^K \alpha_k(x, t, p) m_k.$$

Indeed, this is precisely the relaxation in moments of the global optimization of the Hamiltonian H(x, t, p, u) when the variable u is transformed into the vector m. Thus, every minimizer of the convex formulation (11) attains the minimum value of the non-linear optimal control problem (1)-(2). In the following section we will analyze the connections between the original non-linear, optimal control problem (1)-(2) and its convex relaxation (11).

3 Analysis of the problem

We present here some analytical aspects of the formulation (11) and its relations with the existence of minimizers of the non-linear problem (1)-(2). We face the question of the existence of minimizers from the point of view of absolutely continuous functions [3, 5, 6, 10, 14, 42, 45]. Since relaxation (11) is an optimal control problem with convex structure in the control variable m, we can suppose under mild assumptions that (11) has a minimizer m^* [4, 5, 7, 10, 11, 42]. We will show that the minimizers m^* of the formulation (11) may determine the existence of minimizers of the non-linear optimal control problem (1)-(2). Moreover, we can estimate minimizers of (1)-(2) by calculating the solution of the convex formulation (11). Now we state the first result of our work.

Theorem 6 Let us assume that $u^*(t)$ is a minimizer of the optimal control problem (1)-(2), then the control vector $m^*(t)$ given as:

$$m_k^*(t) = (u^*(t))^k \quad \forall \ k = 0, \dots, K$$
 (12)

is a minimizer of the formulation (11).

Proof. Since $u^*(t)$ is an optimal control for (1) under the form (2), the maximum principle claims that $u^*(t)$ satisfies the global minimization problem:

$$H(x^{*}(t), t, p^{*}(t), u^{*}(t)) = \min_{u \in \mathbb{R}} H(x^{*}(t), t, p^{*}(t), u)$$
(13)

where $x^*(t)$ and $p^*(t)$ satisfy the boundary value problem:

$$\frac{dx}{dt} = g(x, t, u^*(t))$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x}(x, t, p, u^*(t))$$

$$x(0) = x_0, \quad p(1) = \frac{\partial F}{\partial x}\left(x(1)\right).$$
(14)

On the other hand, the Hamiltonian function H has the polynomial form (4) and $u^*(t)$ solves the global minimization problem (13), therefore the vector $m^*(t) \in \mathbb{R}^{k+1}$ given as:

$$m_k^*(t) = (u^*(t))^k \qquad k = 0, \dots, K$$

solves the semidefinite program:

$$\min_{n \in \overline{M}} \quad \widetilde{H}(x^*(t), t, p^*(t), m) \tag{15}$$

where the functions $x^*(t)$, $p^*(t)$ in (15) come from the solution of the boundary value problem (14). Since g and $\frac{\partial H}{\partial x}$ in (14) have a polynomial form in the

variable u^* , and every appearance of the k-th power of $u^*(t)$ can be replaced by $m_k^*(t)$, then the boundary value problem (14) can be expressed as

$$\frac{dx^*}{dt} = \tilde{g}(x^*(t), t, m^*(t))$$

$$\frac{dp^*}{dt} = -\frac{\partial \tilde{H}}{\partial x}(x^*(t), t, p^*(t), m^*(t))$$

$$x(0) = x_0, \quad p(1) = \frac{\partial F}{\partial x}\left(x(1)\right)$$
(16)

where $\tilde{g}(x,t,m) = \sum_{k=0}^{N_2} c_k(x,t)m_k$ and $\frac{\partial \tilde{H}}{\partial x}$ is the formal partial derivative of \tilde{H} with respect to the variable x. As $m^*(t)$ solves (15) and $x^*(t)$, $p^*(t)$ satisfy (16) we have:

$$\widetilde{H}(x^{*}(t), t, p^{*}(t), m^{*}(t)) = \min_{m \in \overline{M}} \quad \widetilde{H}(x^{*}(t), t, p^{*}(t), m)$$
(17)

which is the maximum principle's necessary condition of the convex formulation (11). Since the relaxation (11) is convex, the maximum principle's necessary conditions are also sufficient to guarantee optimality. Thus, $m^*(t)$ is a minimizer of (11).

Corollary 7 If m^* is a minimizer of (11) satisfying (12), $m_1^*(t)$ is a minimizer of (1)-(2).

These results settle a correspondence between the minimizers of (1) and the minimizers of (11) satisfying (12). Thus, problem (1) admits a unique minimizer if and only if its formulation (11) has a unique minimizer satisfying (12). This situation is particularly convenient in order to calculate minimizers of (1), because we only have to solve the convex formulation (11) which is more appropriated to be handled by high performance, non-linear programming techniques.

Corollary 8 If all the minimizers of the formulation (11) do fail in satisfying the expression (12), then the problem (1)-(2) does lack of minimizers.

This result provides a practical method to certify the lack of minimizers in particular non-convex, optimal control problems given in the form (1)-(2). Although it is an expensive computational task, we can determine the lack of minimizers in (1) by checking all the minimizers of (11). We will present some examples in the following section. On the other hand, under particular conditions on the Hamiltonian H we can certify existence of minimizers for (1).

Corollary 9 If the Hamiltonian H(x,t,p,u) is coercive and it has a unique global minimum in u irrespective of the values of x, t and p, then every minimizer of the formulation (11) has the form (12). Hence, (1) has at least a minimizer.

Proof. Let us assume that m^* is a minimizer of (11). According to the maximum principle, m^* satisfies (17) where $x^*(t)$ and $p^*(t)$ satisfy the boundary value problem (16). Since the optimal control vector m^* is composed of the moments of the Dirac measure supported in the singleton $argmin(H(x^*(t), t, p^*(t))), m^*$ has the form (12).

Corollary 10 When H(x, t, p, u) is coercive and strictly convex in u, (1) has a minimizer.

This corollary is well known in control theory, see [4, 5, 10, 11, 14]. We would like to remark that formulation (11) provides additional information of (1) in the sense of Young generalized curves [3, 4, 5, 9, 14]. Let us assume that m^* is a minimizer of (11), therefore m^* satisfy (17) where x^* and p^* in turn satisfy the boundary value problem (16). Then, the entries of m^* are the moments of some probability measure supported in $argmin(H(x^*(t), t, p^*(t))))$, which is a finite set with $K' \leq \frac{K}{2}$ points, i.e.

$$m^{*}(t) = \sum_{i=1}^{K'} \lambda_{i}(t)(1, v_{i}(t), v_{i}^{2}(t), \dots, v_{i}^{K}(t))$$
(18)

where $argmin(H(x^*(t), t, p^*(t))) = \{v_1(t), \ldots, v_{K'}(t)\}$ and $\sum_{i=1}^{K'} \lambda_i(t) = 1$ with $\lambda_i(t) \ge 0$ for $i = 1, \ldots, K'$. As the functions \tilde{g} and $\frac{\partial \tilde{H}}{\partial x}$ in (16) are linear in the variable m, it is easy to see that functions λ and v satisfy the global optimization problem

$$\min_{\lambda \in \Delta_K, v \in \mathbb{R}^K} \overline{H}(x^*(t), t, p^*(t), \lambda, v) = \sum_{i=1}^K \lambda_i H(x^*(t), t, p^*(t), v_i)$$
(19)

where Δ_K is the basic (K-1)-simplex in \mathbb{R}^K . Thus, the functions $\{\lambda_i\}$ and $\{v_i\}$ determine the generalized solution $\sum_{i=1}^{K'} \lambda_i(t)v_i(t)$ of (1) because the set of differential equations in (16) transforms into:

$$\frac{dx^{*}}{dt} = \sum_{i=1}^{K'} \lambda_{i}(t)g(x^{*}(t), t, v_{i}(t))$$

$$\frac{dp^{*}}{dt} = -\sum_{i=1}^{K'} \lambda_{i}(t)\frac{\partial H}{\partial x}(x^{*}(t), t, p^{*}(t), v_{i}(t)).$$
(20)

Summarizing, we propose the formulation (11) as a particular convex relaxation of (1) which can be solved by high-performance, numerical methods for convex mathematical programs. Its solution provides either the minimizers of (1) or its optimal generalized curves. This method allows us to deal with highdemanding, non-linear, non-convex problems arising in several fields of science and technology.

4 Computational Treatment

Now we focus on the computational estimation of the solution of the formulation (11) as a non-linear mathematical program. We take a discrete net of points t_0, \ldots, t_{ν} on the interval time [0, 1], a set of design variables $m(t_r)$ intended to represent the control variables $m \in \mathbb{R}^{K+1}$ and the variables $x(t_r)$ intended to represent the state variables $x \in \mathbb{R}^n$. When the points t_r are uniformly distributed on the interval [0, 1], we obtain the mathematical program:

$$\min_{m,x} \sum_{r=0}^{\nu-1} \int_{rh}^{(r+1)h} \sum_{k=0}^{N_1} a_k m_k(r) dt$$
s.t.
$$\frac{x_r - x_{r-1}}{h} = \sum_{k=0}^{N_2} c_k m_k(r) \quad \forall r = 1, \dots, \nu$$

$$m_0(r) = 1 \quad \forall r = 0, \dots, \nu$$

$$(m_{(i+j)}(r))_{i,j=0}^{K/2} \ge 0 \quad \forall r = 0, \dots, \nu$$
(21)

where h is the uniform distance between the discrete net points. In order to represent the matrix inequality condition as a set of non-linear inequalities, we use the fact that all principal subdeterminants of a positive semidefinite matrix are nonnegative [40]. Then, the matrix inequality condition

$$(m_{i+j}(r))_{i,j=0}^{\frac{K}{2}} \ge 0 \tag{22}$$

is expressed as a set of non-linear inequality constraints:

$$D_{\sigma}(m(r)) \ge 0 \quad \forall \sigma = 1, \dots, \pi, r = 0, \dots, \nu$$

where D_{σ} is the explicit form of every principal subdeterminant of the Hankel matrix in (22) and π is the number of its principal subdeterminants. In this way, we have transformed the optimal control problem (1) into a non-linear, convex, mathematical program. Notice that coefficients a_k and c_k may depend on xand t. In order to solve this kind of high dimensional, non-linear mathematical programs, we use standard professional software based on Quasi-Newton , Least Squares, Gauss-Newton, Sequential Quadratic Programming and Trust Region Methods [34, 41, 43, 44].

4.1 Commented Results

We explain in full detail several examples of non-linear optimal control problems analyzed by the method of moments proposed here.

4.1.1 Example 1

We illustrate the success of the method proposed in this work by solving the following optimal control problem:

$$\min_{u(t)} \int_{0}^{1} (x - 10t)^{2} dt$$

s.t. $\dot{x} = u^{2} - ux + x$
 $x(0) = 0$ (23)

where the cost function is intended to find the best trajectory to approximate the straight line $t \to 10t$ on the interval [0, 1]. Its convex formulation (11) takes the form:

$$\min_{m} \int_{0}^{1} (x - 10t)^{2} dt$$
s.t. $\dot{x} = m_{2} - m_{1}x + x$

$$\begin{bmatrix} 1 & m_{1}(t) \\ m_{1}(t) & m_{2}(t) \end{bmatrix} \ge 0$$
 $\forall t \in [0, 1], \quad x(0) = 0.$
(24)

Next we write down this formulation as the discrete, non-linear mathematical program:

$$\min_{(m,x)} \sum_{r=1}^{\nu} \left[(x_r - 10rh)^2 + (x_{r-1} - 10(r-1)h)^2 \right] \frac{h}{2}$$
s.t. $\frac{x_r - x_{r-1}}{h} = m_2(r) - m_1(r)x_r + x_r$

$$\begin{bmatrix} 1 & m_1(r) \\ m_1(r) & m_2(r) \end{bmatrix} \ge 0 \text{ for } r = 1, \dots, \nu \text{ and } x_0 = 0.$$
(25)

Figure 1(a) compares our result against the trajectory that we should follow. Figure 1(b) shows the optimal control that we obtain from the first moment m_1^* . This is indeed a minimizer as m^* satisfies the condition (12).

4.1.2 Example 2

Here we focus on the optimal control problem

$$\min_{u(t)} \int_{0}^{1} (x - t^{2})^{2} dt$$
s.t. $\dot{x} = u^{2} - ux + x$

$$x(0) = 0$$
(26)

which is intended to approximate the curve $t \to t^2$ under the same, non-linear control system of the Example 1. The results obtained in this case are summarized in Figure 2.



Figure 1: Non-linear control system $\dot{x} = u^2 - ux + x$.



(a) Following the path $t \to t^2$



(b) Optimal control

Figure 2: Non-linear control system $\dot{x} = u^2 - ux + x$.

4.1.3 Example 3

Now we analyze a non-linear, non-convex control system with two state variables and some particular cost function intended to follow the path: $t \to (t, t)$ in the x-y plane.

$$\min_{u(t)} \int_{0}^{1} (x-t)^{2} + (y-t)^{2} dt$$

s.t. $\dot{x} = u$
 $\dot{y} = (1-u^{2})^{2} + x^{2}$
 $x(0) = 0, \ y(0) = 0.$ (27)

We can not apply here numerical integration on the Hamilton equations since the Hamiltonian is a non-linear, non-convex expression of the control variable u. The convex formulation (11) of this problem is:

$$\min_{(m,x)} \int_{0}^{1} (x-t)^{2} + (y-t)^{2} dt$$
s.t. $\dot{x} = m_{1}$
 $\dot{y} = 1 - 2m_{2} + m_{4} + x^{2}$

$$\begin{bmatrix} 1 & m_{1}(t) & m_{2}(t) \\ m_{1}(t) & m_{2}(t) & m_{3}(t) \\ m_{2}(t) & m_{3}(t) & m_{4}(t) \end{bmatrix} \ge 0$$
 $\forall t \in (0,1), \ x(0) = 0, \ y(0) = 0.$
(28)

The corresponding discrete model takes the following form:

$$\min_{(m,x,y)} \sum_{r=1}^{\nu} \left[(x_r - rh)^2 + (y_r - rh)^2 + (x_{r-1} - (r-1)h)^2 + (y_{r-1} - (r-1)h)^2 \right] \frac{h}{2}$$
s.t.
$$\frac{x_r - x_{r-1}}{h} = m_1(r)$$

$$\frac{y_r - y_{r-1}}{h} = 1 - 2m_2(r) + m_4(r) + x_r^2$$

$$\begin{bmatrix} 1 & m_1(r) & m_2(r) \\ m_1(r) & m_2(r) & m_3(r) \\ m_2(r) & m_3(r) & m_4(r) \end{bmatrix} \ge 0 \quad \forall r = 1, \dots, \nu$$
with $x_0 = 0, \ y_0 = 0$
(29)

whose solution for $\nu = 20$ points is given in the Figure 3. It is worth noticing here that we use a cost functional intended to follow the parametrized curve $t \rightarrow (t, t)$ in the (x, y) plane.



(a) two state variables: x and y

(b) Optimal control

Figure 3: Non-linear optimal control problem.

4.1.4 Example 4

We use the same cost function of the previous example, but we introduce here a different non-linear, non-convex control system:

$$\min_{u(t)} \int_0^1 (x-t)^2 + (y-t)^2 dt$$

s.t. $\dot{x} = u$
 $\dot{y} = (1-u)^2 (2-u)^2 + x^2$
 $x(0) = 0, \quad y(0) = 0.$ (30)

By taking $\nu = 20$ points in the discrete model, we obtain the results summarized in Figure 4.



(a) two state variables: \boldsymbol{x} and \boldsymbol{y}

(b) Optimal control

Figure 4: Non-linear, non-convex control system.

4.1.5 Example 5

We present here a particular case taken from control engineering in which we must minimize the steady state error. The corresponding optimal control problem is:

$$\min_{u(t)} \int_{0}^{1} (x-1)^{2} + (y-0.7)^{2} dt$$

s.t. $\dot{x} = u$
 $\dot{y} = u^{2}(1-u^{2})^{2} + x(u^{2} - \frac{1}{2})$
 $x(0) = 0, \ y(0) = 0.$ (31)

Its convex formulation follows:

$$\min_{(m,x,y)} \int_{0}^{1} (x-1)^{2} + (y-0.7)^{2} dt$$
s.t. $\dot{x} = m_{1}$
 $\dot{y} = m_{2} - 2m_{4} + m_{6} + xm_{2} - \frac{x}{2}$

$$\begin{bmatrix} 1 & m_{1}(t) & m_{2}(t) & m_{3}(t) \\ m_{1}(t) & m_{2}(t) & m_{3}(t) & m_{4}(t) \\ m_{2}(t) & m_{3}(t) & m_{4}(t) & m_{5}(t) \\ m_{3}(t) & m_{4}(t) & m_{5}(t) & m_{6}(t) \end{bmatrix} \ge 0$$

$$\forall t \in [0, 1]$$
and $x(0) = 0, \ y(0) = 0$.
$$(32)$$

Its discrete model is:

$$\min_{(m,x,y)} \sum_{r=1}^{\nu} \left((x_r - 1)^2 + (x_{r-1} - 1)^2 + (y_r - 0.7)^2 + (y_{r-1} - 0.7)^2 \right) \frac{h}{2}$$
s.t.
$$\frac{x_r - x_{r-1}}{h} = m_1(r)$$

$$\frac{y_r - y_{r-1}}{h} = m_2(r) - 2m_4(r) + m_6(r) + x_r m_2(r) - \frac{x_r}{2}$$

$$\begin{bmatrix} 1 & m_1(r) & m_2(r) & m_3(r) \\ m_1(r) & m_2(r) & m_3(r) & m_4(r) \\ m_2(r) & m_3(r) & m_4(r) & m_5(r) \\ m_3(r) & m_4(r) & m_5(r) & m_6(r) \end{bmatrix} \ge 0 \quad \forall r = 1, \dots, \nu$$
and $x_0 = 0, y_0 = 0.$
(33)

The results are shown in Figure 5. Certainly, the steady error has been eliminated. See [35, 36, 37].

4.1.6 Example 6

We propose here a non-linear, non-convex Mayer problem:

$$\min_{u} (x(1) - 0.5)^{2} + (y(1) - 0.5)^{2}$$
s.t. $\dot{x} = u$
 $\dot{y} = (1 - u)^{2}(2 - u)^{2} + x^{2}$
 $x(0) = 0, \quad y(0) = 0.$
(34)

m_{1}^{*}	m_2^*	m_3^*	m_4^*	$m_2^* - (m_1^*)^2$	$m_3^* - (m_1^*)^3$	$m_4^* - (m_1^*)^4$
0.01	0.001	0.0001	0.0001	0.00	0.00	0.00
0.5613	0.3175	0.1831	0.1097	0.00	0.01	0.01
0.5685	0.3244	0.1878	0.1108	0.00	0.00	0.01
0.5646	0.3246	0.1947	0.1278	0.01	0.01	0.03
0.5641	0.3238	0.1967	0.1398	0.01	0.02	0.04
0.5641	0.3232	0.1971	0.1471	0.00	0.02	0.05
0.5642	0.3229	0.197	0.1507	0.00	0.02	0.05
0.5643	0.3228	0.1966	0.1513	0.00	0.02	0.05
0.5647	0.3229	0.1959	0.149	0.00	0.02	0.05
0.5653	0.3231	0.1946	0.1438	0.00	0.01	0.04
0.5664	0.3235	0.1925	0.1351	0.00	0.01	0.03
0.5684	0.3244	0.1891	0.121	0.00	0.01	0.02
0.5685	0.3231	0.1854	0.1064	0.00	0.00	0.00
0.5627	0.3166	0.1797	0.1021	0.00	0.00	0.00
0.5542	0.3071	0.1717	0.096	0.00	0.00	0.00
0.5421	0.2939	0.1607	0.0879	0.00	0.00	0.00
0.5254	0.276	0.1464	0.0777	0.00	0.00	0.00
0.5019	0.2518	0.1281	0.0653	0.00	0.00	0.00
0.4753	0.2256	0.1112	0.0537	0.00	0.00	0.00
0.4445	0.1969	0.0941	0.0702	0.00	0.01	0.03

Table 1: Dirac measure moments testing

Its convex formulation is:

$$\begin{array}{ll}
\min_{m,x,y} & (x(1) - 0.5)^2 + (y(1) - 0.5)^2 \\
\text{s.t.} & \dot{x} = m_1 \\
& \dot{y} = 4 - 12m_1 + 13m_2 - 6m_3 + m_4 + x^2 \\
& \begin{bmatrix} 1 & m_1 & m_2 \\ m_1 & m_2 & m_3 \\ m_2 & m_3 & m_4 \end{bmatrix} \ge 0 \\
& x(0) = 0, \quad y(0) = 0.
\end{array}$$
(35)

Figure 6 shows the results obtained after solving the corresponding mathematical program.

m_1	m_2	m_3	m_4
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1
0	1	0	1

Table 2: Table of moments which are not implied by a Dirac measure

4.1.7 Example 7

The following non-convex, optimal control, Mayer problem

$$\min_{u} (x(1) + 3.5)^{2} + (y(1) + 1.75)^{2}$$
s.t. $\dot{x} = \frac{1}{2}u^{2}(u - 4) + \frac{3}{4}(u - 5)$
 $\dot{y} = x$
(36)

$$x(0) = 0, \quad y(0) = 0$$

is intended to find the optimal control which will take the state variables to a precise point in a determined time. Figure 7 shows the results.

4.1.8 Example 8

Consider two coupled water storage tanks whose capacities are C_1 and C_2 respectively. There are two values, one of them has resistivity R_1 and connects both tanks, the other one controls the outward flow of the second tank with resistivity R_2 . The inward flow of water into the first tank is w and it is administered by an electro-value, which in turn is driven by an input voltage u. The



Figure 5: Steady state error minimization





(a) two state variables: x and y

(b) Optimal control

Figure 6: Non-convex Mayer problem

operation of the electro-valve can be represented by the non-linear, non-convex expression:

$$w = (4 - u^2)^2. (37)$$

The dynamics of this system is:

$$\begin{bmatrix} \dot{h}_1 \\ \dot{h}_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{R_1C_1} & \frac{1}{R_1C_1} \\ \frac{1}{R_1C_2} & -\left[\frac{1}{R_1} + \frac{1}{R_2}\right] \frac{1}{C_2} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \begin{bmatrix} 1/C \\ 0 \end{bmatrix} (4-u^2)^2.$$
(38)

In order to minimize the system's control effort and to reach a precise amount of water in a fixed time, we must minimize the expression:

$$\int_0^1 u^2 dt + (h_1(1) - 0.5)^2 + (h_2(1) - 0.2)^2.$$
(39)

Figures 8(a) and 8(b) show the results obtained by the method of moments.



Figure 7: Non-convex Mayer problem.



Figure 8: Non-linear flow control.

4.2 Existence of minimizers and generalized solutions

When $m^*(t)$ is a minimizer of (11), not necessarily the function $m_1(t)^*$ is a minimizer of (1). This is true only when m^* is the vector of moments of a Dirac measure. Thus, in practice, to guarantee that we are obtaining true minimizers of (1), we must check that a particular minimizer of (11) satisfies (12). We have verified this fact for every example presented in this section. Table 1 shows the results of this verification procedure when applied to the results of the Example 6.

When a particular minimizer m^* of (11) does not fulfill with the expression (12), it still gives the generalized minimizers of (1). Consider the classical Bolza

example in Calculus of Variations:

$$\min_{u} \int_{0}^{1} \{\phi(u) + x(t)^{2}\} dt + x(1)^{2}$$
s.t. $\dot{x} = u$

$$x(0) = 0$$
(40)

where $\phi(u) = (1 - u^2)^2$. After solving its convex formulation (11), we obtain the m^* minimizer shown in Table 2. Clearly, these values do not represent the moments of a Dirac measure, so we conclude that problem (40) lacks of minimizers after checking that its relaxed formulation (11) does not have any other minimizers. In addition, we can see that the vector control $m^*(t)$ is composed by the moments of the two-points supported probability measure $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. In this way, we obtain a generalized solution as $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ for every $t \in [0, 1]$.

5 Concluding Remarks

In this work we have proposed a new method for solving explicitly high demanding, non-linear, non-convex optimal control problems given in Lagrange, Mayer or Bolza forms, provided they can be expressed in the general formulation (1)-(2). From a theoretical point of view, we provide necessary and sufficient conditions for the existence of minimizers of (1)-(2) by using particular features of the minimizers of its relaxed, convex formulation (11). Even in the absence of classical minimizers of (1)-(2), the solution of its relaxed formulation (11)provides minimizers of (1)-(2) in the form of Young's generalized curves. This situation is particularly important in applications where practitioners must deal with non-linear, non-convex, optimal control problems for design and simulation. On the other hand, this work also provides numerical schemes to calculate minimizers of non-linear, non-convex, optimal control problems described by polynomials in science and engineering. To attain this task, scientists must use specialized algorithms for convex optimization on high-demanding, discrete models given as non-linear, convex mathematical programs. Finally, we remark that the success of this method is based on the maximum principle of optimal control theory and the global optimization of polynomial expressions carried out by the method of moments.

6 Afterword

The method of moments in global optimization of polynomials has received a good deal of attention from many outstanding researchers working actively in optimization, operations research, computer sciences and control theory. As instance, we would like to mention three recent and relevant events on this topic which reveal how big is the interest on it, how fruitful are its applications and how many people is engaged with it. They were the 8th International Workshop on High Performance Optimization Techniques Optimization and Polynomials, June 23-25, 2004, CWI, Amsterdam, Netherlands, the ADONET Doctoral school Optimization over Polynomials and Semidefinite Programming, September 12-16, 2005, University of Klagenfurt, Austria and SIAM's Conference on Optimization, May 2005, Stockholm, Sweden. We would like to stress the pioneer works on this topic of Jean B. Lasserre, LAAS-CNRS, Toulouse France, Monique Laurent, CWI, Amsterdam, Netherlands and Pablo Parrillo, MIT, Cambridge, U.S.A.. State of the art papers related with our work which will help the reader to grasp the essentials of this new and amazing topic in modern optimization are [17, 18, 19, 20, 21, 22, 53, 54, 55, 56, 57, 58, 59, 60, 61]. This paper has been influenced in many ways for these works, their authors, other colleagues who participated in different ways in these events and all the current research in global optimization and control with moments. We apologize for not having enough space for listing all of them here.

References

- [1] Akhiezer, N. and M. Krein, *Some Questions in the Theory of Moments*, Translations of Mathematical Monographs, vol **2**, AMS, 1962.
- [2] Shohat, J.A. and J.D. Tamarkin, *The Problem of Moments, Mathematical Surveys*, No. 1, AMS, 1943.
- [3] Young, L.C., Lectures on Calculus of Variations and Optimal Control Theory, Saunders, 1969.
- [4] Roubíček, T., Relaxation in Optimization Theory and Variational Calculus, De Gruyter, 1997.
- [5] Cesari, L., Optimization, Theory and Applications, Springer, New York, 1983.
- [6] Zabczyk, J., Mathematical Control Theory: An Introduction, Birkhäuser, Basel, 1992.
- [7] Fattorini, H.O., Infinite Dimensional Optimization Theory and Optimal Control, Cambridge University Press, 1999.
- [8] Egozcue, J., R. Meziat and P. Pedregal, From a nonlinear, non-convex variational problem to a linear convex formulation, J. Applied Mathematics and Optimization, vol 47, pp. 27-44, 2003.
- [9] Pedregal, P., Parametrized Measures and Variational Principles, Birkhauser, 1997.
- [10] Milyutin, A.A. and N.P. Osmolovskii, *Calculus of Variations and Optimal Control*, Transactions of Mathematical Monographs, Vol. 180, AMS, 1998.
- [11] Muñoz, J. and P. Pedregal, A refinement on existence results in nonconvex optimal control, J. Applied Mathematics and Optimization., Vol 41, 129-140.
- [12] Balder, E.J., Young measure techniques for existence of Cournot-Nash-Walras equilibria. Topics in Mathematical Economics and Game Theory (M. Wooders, ed.), Fields Institute Communications 23, pp. 31-39, American Mathematical Society, Providence, 1999.
- [13] Balder, E.J., Lectures on Young measure theory and its applications in economics. (Workshop di Theoria della Misura e Analisi Reale, Grado, 1997), Rend. Istit. Mat. Univ. Trieste.
- [14] Mordukhovich, B.S., Existence theorems in non-convex optimal control, Calculus of Variations and Optimal Control, Technion, Chapman and Hall, 1988.

- [15] Meziat, R. J., El Método de los Momentos para Problemas Variacionales No Convexos, tesis de doctorado, Universidad Politécnica de Cataluña, Barcelona, 2001.
- [16] Krein, M.G. and A.A. Nudelman, *The Markov moment problem and ex*tremal problems, Translations of Mathematical Monographs, vol. 50, AMS, 1977.
- [17] Lasserre, J., Global optimization with polynomials and the problem of moments, SIAM J. Optim., vol 11, 796-817, 2001.
- [18] Lasserre, J., Semidefinite programming vs lp relaxations for polynomial programming, J. of Math. of Operations Research, vol 27, 347-360, 2002.
- [19] Henrion D. and J.B. Lasserre GloptiPoly : Global Optimization over Polynomials with Matlab and SeDuMi., ACM Trans. Math. Soft., 29, 165–194, 2003.
- [20] Lasserre, J. and Prieto-Rumeau T., SDP vs. LP Relaxations for the Moment Approach in Some Performance Evaluation Problems., Stoch. Models, 20, 439–456, 2004.
- [21] Lasserre, J., SOS approximations of polynomials nonnegative on a real algebraic set, SIAM J. Optim., 16, 610–628, 2005.
- [22] Lasserre, J., A sum of squares approximation of nonnegative polynomials, SIAM, J. Optim., 16, 751–765, 2006.
- [23] Ben-Tal, A. and A. Nemirovsky, Lectures on Modern Convex Optimization, SIAM, 2001.
- [24] Meziat, R., Analysis of non-convex polynomial programs by the method of moments, in Frontiers in Global Optimization, C. Floudas and P. Pardalos, eds., Nonconvex Optimization and its Applications Series, vol. 74, p. 353-372, Kluwer, 2003.
- [25] Meziat, R., The method of moments in global optimization, Journal of Mathematical Sciences., vol 116, p. 3303-3324, Kluwer, 2003.
- [26] Meziat, R., Analysis of two dimensional non-convex variational problems, in Optimization and Control with Applications, L. Qi, K. Teo and X. Yang, eds., Springer, 2005.
- [27] Nesterov, Y., Squared functional systems and optimization problems in High Performance Optimization, Frenk, H., K. Roos and T. Terlaky, eds., Kluwer, 2000.
- [28] Nesterov Y. and A. Nemirovsky, Interior-Point Polynomial Methods in Convex Programming, Studies in Applied Mathematics, Society for Industrial and Applied Mathematics, vol. 13, 1994.

- [29] Curto, R. and L.A. Fialkow, Recursiveness, positivity and truncated moment problems, Houston Journal of Mathematics, Vol 17, p. 603-635, 1991.
- [30] Curto, R. and L.A. Fialkow, Solution of the truncated complex moment problem for flat data, Memoirs of the AMS, No. 568, 1996.
- [31] Curto, R. and L.A. Fialkow, *The truncated complex k-moment problem*, Transactions of the AMS, vol 352, No. 6, p. 2825-2855, 2000.
- [32] Castillo, E. et al., Building and Solving Mathematical Programming Models in Engineering and Science, Wiley, New York, 2000.
- [33] Kirk, D.E., Optimal Control Theory: An Introduction, Prentice-Hall, 1970.
- [34] The MathWorks Inc., Optimization Toolbox For Use with Matlab. User's Guide, 2000.
- [35] Ogata, K., Discrete-Time Control Systems, Prentice-Hall, 1997.
- [36] Ogata, K., Modern Control Engineering, Prentice-Hall, 2001.
- [37] Zhou, K., J.C. Doyle and K. Glover, *Robust and Optimal Control*, Prentice Hall, 1996.
- [38] Boyd, S. and L. Vandenberge, *Convex Optimization*, Cambridge University Press, 2004.
- [39] Nocedal, J. and S. Wright, Numerical Optimization, Springer-Verlag, 1999.
- [40] Horn, A. and R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [41] Fletcher, R., Practical methods of optimization, John Wiley & Sons, 1987.
- [42] Pedregal, P., Introduction to Optimization, Springer Verlag, 2003.
- [43] Conn, A.R., N. Gould and P. Toint, Trust-Region Methods, SIAM, 2000.
- [44] Luenberger, D. E., Linear and Nonlinear Programming, Addison-Wesley, 1984.
- [45] Craven, B.D., Control and Optimization, Chapmann & Hall, 1995.
- [46] Betts, J.T., Practical Methods for Optimal Control Using Non-linear Programming, SIAM, 2001.
- [47] Sussmann, H.J., Geometry and optimal control, in Mathematical Control Theory, J. Baillieul and J.C. Willems, eds., Springer, 1999.
- [48] Jurdjevic, V., Optimal control, geometry, and mechanics, in Mathematical Control Theory, J. Baillieul and J.C. Willems, eds., Springer, 1999.

- [49] Bloch, A.M. and P.E. Crouch, Optimal control, optimization and analytical mechanics in Mathematical Control Theory, J. Baillieul and J.C. Willems, eds., Springer, 1999.
- [50] Haslinger J. and R.A.E. Mäkinen, Introduction to shape optimization : theory, approximation, and computation, SIAM, 2003.
- [51] Roubíček, T. and Valášek, M., T., Optimal control of causal differential algebraic systems, J. of Math. Anal. Appl., vol. 269, No. 2, Academic Press, 2002.
- [52] Folcher, J.P. and C. Andriot, *Multiobjective control for robot telemanipulators*, in Advances in Linear Matrix Inequality Methods in Control, L.E. Ghaoui and S. Niculescu, SIAM, 2000.
- [53] Laurent, M., Revisiting two theorems of Curto and Fialkow on moment matrices, Proceedings of the AMS, 133, no. 10, 2965–2976, 2005.
- [54] Laurent, M. and D. Jibetean, Semidefinite approximations for global unconstrained polynomial optimization, SIAM Journal on Optimization, 16(2), 490–514, 2005.
- [55] Henrion, Didier; Garulli, Andrea (Eds.), Positive Polynomials in Control Series, Lecture Notes in Control and Information Sciences, Vol. 312, 2005.
- [56] Parrilo, P. and B. Sturmfels, *Minimizing polynomial functions*, in Algorithmic and quantitative real algebraic geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, AMS., Vol. 60, 83–99, 2001.
- [57] Parrilo, P. and S. Lall, Semidefinite programming relaxations and algebraic optimization in Control, European Journal of Control, Vol. 9, No. 2-3, 2003.
- [58] Parrilo, P., Semidefinite programming relaxations for semialgebraic problems, Mathematical Programming Ser. B, Vol. 96, No.2, 293–320, 2003.
- [59] Prajna, S., Parrilo P. and A. Rantzer, Nonlinear control synthesis by convex optimization, IEEE Transactions on Automatic Control, Vol. 49, No. 2, 2004.
- [60] Mazzaro, M.C., P.A. Parrilo and R.S. Sánchez Peña, Robust identification toolbox, Latin American Applied Research, Vol. 34, No. 2, 91–100, 2004.
- [61] K. Gatermann and P.A. Parrilo, Symmetry groups, semidefinite programs, and sums of squares, Journal of Pure and Appl. Algebra, Vol. 192, No. 1-3, 95–128, 2004.
- [62] Meziat, R. and J. Villalobos, Analysis of microstructures and phase transitions in one dimensional elasticity by convex optimization, submitted to Journal of Strutural and Interdisciplinary Optimization.

Acknowledgments

This work was supported by the following research grants:

- 1. Proyecto Semilla, Fondo de Investigaciones, Facultad de Ciencias de la Universidad de los Andes, 2003.
- Optimización Global de Problemas Variacionales, Proyecto del Plan Nacional de Ciencias Básicas, COLCIENCIAS 1204-05-13627, 2004.
- Aplicaciones Avanzadas de las Medidas de Young, Proyecto MTM-2004-07114, Spain.
- 4. Proyecto conjunto Aula Cimne-Uniandes, 2003-2008.
- 5. Proyecto Jóvenes Investigadores, Diego Patiño, Colciencias, 2005.
- 6. Matching Grant LMGP-0311-1110522121, Sun Microsystems, 2005.