On inverse optimal control via polynomial optimization

Jean-Bernard Lasserre, Jérémy Rouot

On inverse optimal control via polynomial optimization

Jérémy Rouot and Jean-Bernard Lasserre

***Abstract*— We consider the class of control systems where the differential equation, state and control system are described by polynomials. Given a set of trajectories and a class of Lagrangians, we are interested to find a Lagrangian in this class for which these trajectories are optimal. To model this inverse problem we use a relaxed version of Hamilton-Jacobi-Bellman optimality conditions, in the continuity of previous work in this vein. Then we provide a general numerical scheme based on polynomial optimization and positivity certificates, and illustrate the concepts on a few academic examples.**

I. INTRODUCTION

Given a system dynamics, control and/or state constraints, and a database $\mathcal{D}$ of feasible state-control trajectories $\{x(t;x_0,t_0),u(t;x_0,t_0)\}$ from several initial states (and different initial times) (x_0,t_0), the *Inverse Optimal Control Problem* (IOCP) consists of computing a Lagrangian L such that each trajectory of the database $\mathcal{D}$ is an optimal solution of the (direct) optimal control problem (OCP) for the cost functional $\int_{t_0}^{T} L(x(t),u(t))dt$ to minimize. Even though it is an interesting problem on its own, it is also of primary importance in Humanoid Robotics where one tries to understand human locomotion. In fact whether or not human locomotion obeys some optimization principle is an open issue, discussed for instance in several contributions in [13].

As most interesting inverse problems, the (IOCP) is in general an ill-conditioned and ill-posed problem. For instance, for a source of ill-posedness, consider an (OCP) with terminal state constraint $x(T)\in X_T$ and with Lagrangian L. Let $\phi : X\times[0,T]\to\mathbb{R}$ be its associated optimal value function and let $g : X\to\mathbb{R}$ be a continuously differentiable function such that $g(x)=0$ for all $x\in X_T$. If the couple $(x^\star(t;x_0,t_0),u^\star(t;x_0,t_0))$ is an optimal solution of the (OCP) then it is also an optimal solution of the (OCP) with new Lagrangian $L-\langle\nabla g,f\rangle$ (where f is the vector field of the dynamics) and the optimal value function associated with the new (OCP) is $\phi+g$. See e.g. [18] for a discussion. Therefore in full generality, recovering a Lagrangian that has an intrinsic physical meaning might well be searching for a needle in a haystack. This is perhaps why previous works have considered the (IOCP) with restrictive assumptions on the class of Lagrangian to recover. Some restrictions are also more technical and motivated by simplifying the search process.

In the recent work [18] the authors have proposed to consider the (IOCP) in relatively general framework. In this framework an *ε-optimality certificate* for the trajectories in the database $\mathcal{D}$ translates into natural "positivity constraints" on the unknown optimal value function ϕ and Lagrangian L, in a *relaxed version* of the Hamilton-Jacobi-Bellman (HJB) equations. Moreover some natural normalizing constraint allows to avoid the source of ill-posedness alluded to above. Then the (IOCP) is posed as a *polynomial optimization* problem and solved via a hierarchy of semidefinite relaxations indexed by $d\in\mathbb{N}$, where the unknown is a couple (ϕ,L) of polynomials of potentially large degree d; for more details the reader is referred to [18].

In the cost criterion $\min\varepsilon+\gamma\|L\|_1$, the regularizing parameter $\gamma>0$ controls the tradeoff between minimizing ε and minimizing the sparsity-inducing ℓ_1-norm $\|L\|_1$ of the Lagrangian. (One assumes that a Lagrangian L with a physical meaning will have a sparse polynomial approximation.) In some examples treated in [18] one clearly sees the importance of γ to recover a sparse meaningful Lagrangian; the larger is γ the sparser is the Lagrangian (but at the cost of a larger ε, due to a limited computing capacity).

A. Contribution

As in [18] the approach in this paper is based on the relaxed-HJB framework to state an ε-optimality certificate for the trajectories in the database $\mathcal{D}$. Indeed as already mentioned in [18], HJB-optimality equations is the "perfect tool" to certify global optimality of a given trajectory (Bellman optimality principle is sometimes called a *verification* theorem). On the other hand, instead of searching for L as a polynomial of potentially large degree as in [18], we now restrict the search of L to some *dictionary* which is a family of polynomials of low degree (for instance convex quadratic). In this respect we also follow other previous approaches (e.g. [15], [17]) which impose restrictive assumptions on the Lagrangian. A normalizing condition on L avoids some previously mentioned ill-posedness and in principle no regularizing parameter γ is needed (that is, one only minimizes ε which appears in the ε-optimality certificate). Finally, and again as in [18], the (IOCP) is solved by a hierarchy of semidefinite relaxations. The quality of the solution improves in the hierarchy but of course at a higher computational cost.

To compare this approach with that of [18] we have considered four optimal control problems (of modest size): The LQG, the Brockett double integrator, and two minimum-time problems. In all cases we obtain better results as the Lagrangian L is recovered exactly at a semidefinite relaxation of relatively low order the hierarchy. In particular no regularizing parameter γ is needed. At last but not least, a by-product of the approach is to also provide a means to

J. Rouot is with LAAS-CNRS, 7 avenue du colonel Roche, F-31400 Toulouse, France, j.rouot@laas.fr

Jean B. Lasserre is with LAAS-CNRS and Institute of Mathematics University of Toulouse LAAS, 7 avenue du Colonel Roche 31077 Toulouse, France. lasserre@laas.fr

detect whether the *dictionary* of possible Lagrangian is large enough. Indeed if the resulting optimal value $\varepsilon^\star$ remains relatively *high* after several semidefinite relaxations, one may probably infer that the dictionary is too small to contain a good Lagrangian.

In section II we define the (IOCP) while the relaxed Hamilton-Jacobi-Bellman optimality conditions are presented in Section III (with some discussion on the ill-posedness of the problem). The numerical scheme consisting of a hierarchy of semidefinite relaxations is described in Section IV. The method is illustrated on four academic examples and compared with [18], before some discussion of the relative merits of the method.

II. Setting up the problem

A. Context

Let n and m be nonzero integers. Consider the control system

$$\dot{x}(t) = f(x(t), u(t)) \tag{1}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m$ is a (smooth) vector field, and $u : [t_0, T] \to U$ (compact) $\subset \mathbb{R}^m$ are the controls. A control $u(\cdot)$ is said admissible on $[t_0, T]$ ($T > t_0 > 0$) if the solution $x(\cdot)$ of (1) satisfies $x(t_0) = x_0$ for some $x_0 \in \mathbb{R}^n$ and

$$(x(t), u(t)) \in X \times U \text{ a.e. on } [t_0, T], \quad x(T) \in X_T \tag{2}$$

where X, X_T are compact subsets of $\mathbb{R}^n$. The set of admissible controls is denoted as $\mathcal{U}$. Let us denote by $(\tau, x(\cdot), u(\cdot))$ a feasible trajectory for (1)-(2) which starts at time $\tau \in [0, T]$ in state $x(\tau)$. For an admissible control $u(\cdot)$, we consider only integral costs associated with the corresponding trajectory $x(\cdot)$, and defined by :

$$J(t_0, T, x_0, u(\cdot), L) = \int_{t_0}^{T} L(x(t), u(t))\, \mathrm{d}t, \tag{3}$$

where $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a continuous Lagrangian.

a) Direct optimal control problem (OCP): Given a Lagrangian L and an initial condition $x_0 \in \mathbb{R}^n$, the (OCP) consists in determining a state-control trajectory $(t_0, x(\cdot), u(\cdot))$ solution of (1) with $x(t_0) = x_0$, satisfying the state and control constraints (2) and minimizing the cost (3). We will refer to this problem as $OCP(t_0, x_0, L)$.
The value function $J_L^\star$ associated with $OCP(t_0, x_0, L)$ is defined by

$$J_L^\star(t_0, x_0) = \inf_{u(\cdot) \in \mathcal{U}} J(t_0, T, x_0, u(\cdot), L) \tag{4}$$

if the final time T is fixed, and,

$$J_L^\star(x_0) = \inf_{T, u(\cdot) \in \mathcal{U}} J(0, T, x_0, u(\cdot), L) \tag{5}$$

if the final time T is free.

b) Inverse optimal control problem (IOCP): Given a set $\mathcal{D}$ of controls and their associated feasible trajectories (obtained by observations) satisfying (2), a couple defined by a control system (1) and a class (or dictionary) $\mathcal{L}$ of Lagrangians (each $L \in \mathcal{L}$ defines a cost functional (3)), the (IOCP) consists in finding a Lagrangian $L \in \mathcal{L}$ such that the trajectories of $\mathcal{D}$ are optimal solutions of the (OCP) associated with L.
More precisely, we consider a finite family $\mathcal{D} = \{(t_i, x_i(\cdot), u_i(\cdot))\}_{i \in I}$ of admissible trajectories, indexed by some set I, where $t_i \in [0, T]$ and $u(\cdot) \in \mathcal{U}$ for all $i \in I$. We want to compute $L \in \mathcal{L}$ such that $x_i(\cdot)$ (resp. $u_i(\cdot)$) are optimal state-trajectories (resp. control-trajectories) of $OCP(t_i, x_i(t_i), L), \forall i \in I$.

III. Hamilton-Jacobi-Bellman to certify global-optimality

A. Hamilton-Jacobi-Bellman

Hamilton-Jacobi-Bellman (HJB) equation is the "ideal" tool to provide sufficient (and in a certain sense almost necessary) conditions for a trajectory to be globally-optimal, see [4], [8], [3]. In general, the *value function* is nonsmooth and cannot be identified to a classical solution of the HJB equation. Yet, there are many approaches to characterize the value function from the HJB equation, among others the concept of viscosity solutions [3] or the concept of proximal subgradient [21]. In what follows, a *relaxation* of the (HJB) equation provides a simple certificate of global-optimality for (OCP) of the form $OCP(t_0, x_0, L)$.

Definition 3.1: Let $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and let $\varphi : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable. Given a vector field f, we define the linear operator $\mathcal{H}_f$ which acts on (L, φ) as

$$\mathcal{H}_f(L, \varphi) = L + \frac{\partial \varphi}{\partial t} + \nabla_x \varphi^\intercal f$$

The following certificate was at the core of the method proposed in [18].

Proposition 3.2: Let $\bar{u}(\cdot) \in \mathcal{U}$ be an admissible control associated with a trajectory $\bar{x}(\cdot)$ starting from x_0 at time $t = t_0$. Assume there exists a function $\varphi : [t_0, T] \times X \to \mathbb{R}$, continuously differentiable and such that

$$\mathcal{H}_f(L, \varphi)(t, x, u) \geq 0, \ \forall (t, x, u) \in [t_0, T] \times X \times U, \tag{6}$$

$$\varphi(T, x) = 0, \ \forall x \in X_T, \tag{7}$$

$$\mathcal{H}_f(L, \varphi)(t, \bar{x}(t), \bar{u}(t)) = 0, \ \forall t \in [t_0, T]. \tag{8}$$

Then $\bar{u}(\cdot)$ is an optimal control associated with the trajectory $\bar{x}(\cdot)$ for the problem $OCP(t_0, x_0, L)$.

Proof: For any admissible control $u(\cdot)$ associated with a trajectory $x(\cdot)$, integrating (6) yields $\int_{t_0}^{T} L(x(s), u(s)) \mathrm{d}s + \int_{t_0}^{T} \mathrm{d}\varphi\,(s, x(s)) \geq 0$. Using (7), we obtain $\varphi(t_0, x_0) \leq J(t_0, T, x_0, u(\cdot), L) + \varphi(T, x(T)) = J(t_0, T, x_0, u(\cdot), L)$. The same applied to $\bar{x}(\cdot)$ and $\bar{u}(\cdot)$ yields $\varphi(t_0, x_0) = J(t_0, T, x_0, \bar{u}(\cdot), L)$ and therefore $\bar{u}(\cdot)$ is optimal for $OCP(t_0, x_0, L)$. ∎

B. ε-optimal solution for the inverse problem

Next, we describe slightly weaker conditions than those of Proposition 3.2. They are the basis of our numerical method.

Proposition 3.3: Let $\bar{u}(\cdot)$ be an admissible control associated with a trajectory $\bar{x}(\cdot)$ starting from x_0 at $t = t_0$. Suppose that there exist a real ε and a continuous function $L : X \times U \to \mathbb{R}$ and continuous differentiable function $\varphi : [0,T] \times X \to \mathbb{R}$ such that

$$\mathcal{H}_f(L,\varphi)(t,x,u) \geq 0,\ \forall (t,x,u) \in [t_0,T] \times X \times U, \quad (9)$$

$$-\varepsilon \leq \varphi(T,x) \leq 0,\ \forall x \in X_T, \quad (10)$$

$$\int_{t_0}^{T} \mathcal{H}_f(L,\varphi)(s,\bar{x}(s),\bar{u}(s))\mathrm{d}s \leq \varepsilon. \quad (11)$$

Then $\bar{u}(\cdot)$ is 2ε-optimal for $OCP(t_0,x_0,L)$.

Proof: Following the proof of Proposition 3.2, we get $\varphi(t_0,x_0) \leq J(t_0,T,x_0,u(\cdot),L)$, $\forall u(\cdot) \in \mathcal{U}$ and from condition (11), we get $\varphi(t_0,x_0) \geq J(t_0,T,x_0,\bar{u}(\cdot),L) + \varphi(T,\bar{x}(T)) - \varepsilon$ and therefore $J(t_0,T,x_0,\bar{u}(\cdot),L) \leq \varphi(t_0,x_0) + 2\varepsilon$. ∎

In general the value function is nonsmooth, but in case where it is at least continuous, by the Stone-Weierstrass theorem, it can be approximated by polynomials, uniformly on compact sets, and this approximation will provide lower and upper bounds for the value function. Note that in this case, we can approximate the optimal value as closely as desired.

C. Non well-posedness of the inverse problem

Proposition 3.3 provides us with a simple and powerful tool to certify global-suboptimality of a given set of trajectories, e.g. those given in the database $\mathcal{D}$. However as we have already mentioned, there are many couples (φ, L) that can satisfy (9)-(11), and which make the (IOCP) ill-posed in general. For instance:

- *Trivial solution.* The zero Lagrangian and the zero function satisfy conditions (9)-(10)-(11) with $\varepsilon = 0$, independently of the input trajectories.
- *Conserved quantity.* Let $u(\cdot)$ be an optimal control associated with a trajectory $x(\cdot)$. Suppose that the system dynamics is such that all feasible trajectories satisfy an equation of the form $g(x(t),u(t)) = 0$, $\forall t \in [0,T]$ for some continuous function g. Then $L = g^2$ is a candidate for the inverse problem associated with the zero value function φ.
  Such class of optimal control problems is well spread in applications, especially nonholonomic systems [6], [7]. The latter falls into sub-Riemannian geometry framework and have many applications namely in robotics [5], [12], [2]. They can be formulated as
  $$\begin{aligned} &L(x) = (\dot{x}\cdot\dot{x})^{1/2}, \quad (T \text{ is fixed}) \\ &f(x,u) = \sum_{i=1}^{m} u_i F_i(x),\ x \in X \subset \mathbb{R}^n,\ u \in U \subset \mathbb{R}^m \\ &X_T = \{x_T\},\ x_T \in \mathbb{R}^n \end{aligned} \quad (12)$$
  where $(F_1, \dots F_m)$ are m (smooth) orthonormal vector fields. The Lagrangian $L(x)$ stands for the length of the admissible curve $x(\cdot)$.
  It is well known [6] that this is a length minimization problem equivalent to the energy minimization problem where the Lagrangian is $L(x,u) = \dot{x}\cdot\dot{x}$ and also equivalent to the time minimization problem with $L(x,u) = 1$ and where the control domain is restricted to: $\|u\|_2^2 = 1$.
- *Total variations.* Consider a continuous function $\psi : [0,T] \times X \to \mathbb{R}$ such that ψ vanishes on X_T. Then (φ, L) with $\varphi = -\psi$ and $L = -\nabla_x\psi \cdot f$ is solution to the (IOCP) regardless of the database $\mathcal{D}$!
- *Conic convexity.* Let (L,φ) be a solution of the inverse problem $(IOCP)$ and let $\psi : [0,T] \times X \to \mathbb{R}$ be such that $\psi(T,x) = 0$, $\forall x \in X_T$. If $\tilde{\varphi} = \varphi + \psi$ then the pair $(L + (\nabla_x\psi)\cdot f), \tilde{\phi})$ satisfies the same conditions (9)-(10)-(11) with the same ε and is solution of the same inverse problem.

To illustrate this last issue, we consider the following problem. Let φ_L be a solution of HJB equations associated with the following optimal control problem

$$(P)_L \qquad \begin{aligned} &\min_{T>0, \|u\|_2 \leq 1} \int_0^T L(x(t))\,\mathrm{d}t, \qquad \text{(T is free)} \\ &\dot{x} = u, \\ &x(0) \in X = B_2, \quad x(T) \in X_T = \partial X \end{aligned}$$

where B_2 denotes the unit ball of $\mathbb{R}^2$.

Lemma 3.4: The optimal control law of $(P)_L$ is given by $u = -\nabla\varphi_L / \|\nabla\varphi_L\|_2$ and the Lagrangian L satisfies $L = \|\nabla\varphi_L\|_2$.

Proposition 3.5: Let $p \in \mathbb{N}$ and define $L_p(x) = \|x\|_2^p$. The value function associated with $(P)_{L_p}$ is $J^\star_{L_p}(x) = 1/(p+1)\,(1 - \|x\|_2^{p+1})$. Besides, for all $p \in \mathbb{N}$, problems $(P)_{L_p}$ have the same optimal control law given by $u^\star(x) = x/\|x\|_2$.

For $p \in \mathbb{N}$, let $g_p : X \to \mathbb{R}$ be defined by $g_p(x) = \|x\|_2 - 1/(p+1)\,(\|x\|_2^{p+1} + p)$. Then, $L_p(x) = L_0(x) - \langle \nabla g_p(x), u^\star(x)\rangle$ and $J^\star_{L_p}(x) = J^\star_{L_0}(x) + g_p(x)$.

Normalization.

Due to the convexity of the conditions (9)-(10)-(11), the set of solutions of the inverse problem is a convex cone. Hence, to overcome the non uniqueness of solutions of the inverse problem and to avoid solutions which do not have a physical interpretation, a natural idea is to impose a normalizing constraint (see e.g. [2], [15] or more recently [18]). For instance for free terminal time problems, in [18] one imposes the constraint on (φ, L):

$$|\,1 - \mathcal{A}(L,\varphi)\,| = \varepsilon,$$

where $\mathcal{A}$ is the linear operator:

$$\mathcal{A}(L,\varphi) = \int_{\tilde{X}\times U} \Big[L + \frac{\partial\varphi}{\partial t} + \nabla_x\varphi\cdot f\Big]\mathrm{d}x\mathrm{d}u.$$

This allows to avoid the trivial solution $(\varphi, L) = (0,0)$.

In this paper, we will not use such a normalization. Rather we propose (a) to restrict the search of Lagrangian solutions

L to a class of functions $\mathcal{L}$ introduced in the next section, and (b) include a normalization constraint on L which is more appropriate for our numerical scheme.

IV. Numerical implementation

A. Polynomial and semi-algebraic context

Without any further assumptions on the data X, X_T, U and f, the "positivity constraints" (9)-(10)-(11) cannot be implemented efficiently. However if X, X_T, U are compact basic semi-algebraic sets and f is a polynomial, then we may and will use powerful positivity certificate from real algebraic geometry, e.g. Putinar's Positivstellensatz [20]. In doing so, the positivity constraints (9)-(10)-(11) translate into linear matrix inequalities (LMI). The maximum size of the latter depends on the degree allowed for the *sums-of-squares* (SOS) polynomials involved in the positivity certificate. Of course the higher is the degree, the more powerful is the certificate.

In the sequel the class of Lagrangian $\mathcal{L}$, denoted by $\mathcal{L}_{a,b}$, is a set of polynomials. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_p) \in \mathbb{N}^p$, $|\alpha| = \sum_{i=1}^p \alpha_i$ and for $y = (y_1, \ldots, y_p)$, the notation y^α stands for the monomial $y_1^{\alpha_1} \ldots y_p^{\alpha_p}$. For two vectors $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $u = (u_1, \ldots, u_m) \in \mathbb{R}^m$, and two integers a, b, let $\mathcal{L}_{a,b} \subset \mathbb{R}[x, u]$ be defined by:

$$\mathcal{L}_{a,b} = \Big\{ L \in \mathbb{R}[x,u] \mid L(x,u) = m_a(x)^\intercal\, C_a^x\, m_a(x) + m_b(u)^\intercal\, C_b^u\, m_b(u),\ C_a^x \in \mathbb{S}_+^j,\ j = \binom{n+a}{a},\ C_b^u \in \mathbb{S}_+^k,\ k = \binom{m+b}{b} \Big\}, \tag{13}$$

where $m_d(z)$ denotes the vector of monomials $(z^\alpha)_{|\alpha| \le d}$ which forms a canonical basis of $\mathbb{R}_d[z]$ and $\mathbb{S}_+^d$ is the set of real symmetric and positive semidefinite $d \times d$ matrices.

Remark 4.1: A polynomial of $\mathcal{L}_{a,b}$ is the sum of a polynomial $L_x \in \mathbb{R}_{2a}[x]$ and a polynomial $L_u \in \mathbb{R}_{2b}[u]$ such that L_x and L_u are SOS. Conversely, every SOS polynomial q of $\mathbb{R}_{2d}[x]$ can be written as $q = m_d(x)^\intercal\, Q\, m_d(x)$, $Q \in \mathbb{S}_+$.

To avoid the trivial Lagrangian $L = 0$, we include the normalizing constraint

$$\mathrm{tr}(C_a^x) + \mathrm{tr}(C_b^u) = C,\ C \text{ is a constant.} \tag{14}$$

where $\mathrm{tr}(A)$ denotes the trace of a matrix A, and C_a^x, C_b^u are the matrices which define L in its definition (13). Note that the map $P \in S_+ \mapsto \mathrm{tr}(P) \in \mathbb{R}_+$ is linear and define a norm.

B. Inverse problem formulation

For practical computation, the input data consists of N trajectories $\mathcal{D} = \{(t_i, x_i(\cdot), u_i(\cdot))\}_{i=1,\ldots,N}$ $(t_i, x_i(\cdot), u_i(\cdot))$ starting at time t_i in state $x_i(t_i)$ and admissible for (1)-(2), for all $i = 1, \ldots, N$. In practice, each trajectory $(t_i, x_i(\cdot), u_i(\cdot))$ is sampled at some points (kT/s), $k = 0, 1, \ldots s$, of the interval $[0, T]$, for some fixed integer $s \in \mathbb{N}$. Then with $t_i := k_i T/s$ for some integer $k_i < s$, the database $\mathcal{D}$ is the collection $((k_i T/s, x_i(kT/s), u_i(kT/s))$, $k = k_i, \ldots, s$, $i = 1, \ldots, N$,

Then condition (11) is replaced by :

$$\sum_{k=k_i}^{s} \mathcal{H}_f(L,\varphi)(k_i, x_i(kT/s), u_i(kT/s)) \le \varepsilon,$$

for all $i = 1, \ldots, N$. Alternatively one my also consider

$$\sum_{i=1}^{N} \sum_{k=k_i}^{s} \mathcal{H}_f(L,\varphi)(k_i, x_i(kT/s), u_i(kT/s)) \le \varepsilon.$$

Then consider the following hierarchy of optimization problems indexed by $p \in \mathbb{N}$:

$$\begin{aligned} \varepsilon_p^\star = \inf_{L,\varphi,\varepsilon}\ & \varepsilon \\ \text{s.t. }\ & \mathcal{H}_f(L,\varphi)(t,x,u) \ge 0,\ \forall (t,x,u) \in [0,T] \times X \times U, \\ & \varphi(T,x) \le 0,\ \forall x \in X_T, \\ & \varphi(T,x_k) \ge -\varepsilon,\ k = k_i, \ldots, s, \\ & \sum_{i=1}^{N} \sum_{k=k_i}^{s} \mathcal{H}_f(L,\varphi)(k_i, x_i(kT/s), u_i(kT/s)) \le \varepsilon, \\ & L \in \mathcal{L}_{a,b} \text{ with } \mathrm{tr}(C_a^x) + \mathrm{tr}(C_b^u) = 1, \end{aligned} \tag{$IOCP_{a,b}$}$$

where $\varphi \in \mathbb{R}_{2p}[t,x]$ and $\varepsilon \in \mathbb{R}$.

Remark 4.2: $(IOCP_{a,b})$ is written for control problems with fixed terminal time T. For free terminal time problems the function φ doesn't depend on t.

In $(IOCP_{a,b})$, each positivity constraint on (L, φ) is replaced by a Putinar's positivity certificate [20]. The latter asserts that if a polynomial p is positive on a compact basic semi-algebraic set $K = \{x \mid g_i(x) \ge 0,\ g_i \in \mathbb{R}[x],\ i = 1, \ldots, m\}$ then p can be written as

$$p(x) = \sum_{i=0}^{m} g_i(x) \sigma_i(x), \quad \forall x \in \mathbb{R}^n, \tag{15}$$

for some SOS polynomials σ_i, $i = 0, \ldots, m$ (and where $g_0(x) := 1$ for all x). In addition we impose the degree bound $\deg(\sigma_i\, g_i) \le 2p$, $i = 0, \ldots, m$.

In doing so $(IOCP_{a,b})$ becomes a hierarchy of *semidefinite programs* (SDP) indexed by $p \in \mathbb{N}$ [10], [9]. The size of each SDP in the hierarchy depends on p, i.e., on the strength of Putinar's positivity certificate (15) used in $(IOCP_{a,b})$. In practice this size is limited by the capabilities of semidefinite solvers.

Our implementation is based on the `YALMIP` toolbox [14] which provides a SOS interface to handle polynomial constraints and then solves the resulting SDP problem by running the `MOSEK` solver [16]. We can handle problems up to $6, 7$ variables (t, x, u) and degree 12 for φ. For larger size problems, heuristics and/or techniques exploiting the structure of the problem should be used. Also, using the recent alternative positivity certificates (to (15)) proposed in [11] and [1] should be helpful.

V. Illustrative examples

We have tested our numerical method described in Section IV on four examples also analyzed in [18].

A. Settings

The input of our numerical scheme consists of:

- final time T if the problem has fixed finite horizon,
- state constraint set X and the control set U,
- state (and final state) constraints X, X_T,
- vector field f,
- class of polynomials $\mathcal{L}_{a,b}$ (13) to which we restrict the search for the Lagrangian L.

Since for problems considered here, at each point $(t,x) \in [0,T] \times X$ we know the associated optimal control $u(t)$, the database $\mathcal{D}$ now consists of the collection $((t_k, x_k, u(t_k))_{k=1,\ldots,N} \in [0,T] \times X \times U$ where for each k the couple (t_k, x_k) is randomly generated in some set $S_{\mathcal{D}} \subset [0,T] \times X$.

The output of (IOCP) is a triple $(L, \varphi, \varepsilon^\star)$ where L is a polynomial Lagrangian, φ is a polynomial value function and $\varepsilon^\star$ is a parameter which quantify the sub-optimality of the solutions associated with L. More precisely, according to Proposition (3.3), if $J^\star_L(t_0, x_0)$ is the value function of $OCP(t_0, x_0, L)$ then

$$\begin{aligned}
&\varphi(t_k, x_k) \geq J^\star_L(t_k, x_k) - 2\,\varepsilon^\star,\ k = 1, \ldots, N,\\
&\varphi(T, x_k) \geq -\varepsilon^\star,\ k = 1, \ldots, N,\\
&\varphi(t,x) \leq J^\star_L(t,x), \quad \varphi(T,x) \leq 0,\ \forall (t,x) \in [0,T] \times X.
\end{aligned}$$

B. Benchmark direct problems

We present the optimal control problems that we considered for our numerical computation. In all cases we know the value function $J^\star_{\bar{L}}$ associated with the Lagrangian $\bar{L}$.

1) Linear Quadratic case:

$$T = 1,\ X = B_2,\ X_T = \mathbb{R}^2,\ U = \mathbb{R}^2, f = (x_2, u)^\intercal.$$

where B_2 denotes the unit ball in $\mathbb{R}^2$. This is the celebrated LQ-problem and here the Lagrangian to recover is $\bar{L} = 2x_1^2 + 1/2\, x_1 x_2 + x_2^2 + u^2$.
The optimal control law is given by the closed-loop control $u(t) = I_2\, B\, E(t) x(t)$ where, for this problem $B = (1 \quad 0)$, $E(.)$ is solution of the corresponding Riccati equation ($E(T) = 0$) and the value function is given by $J^\star_{\bar{L}}(t, x(t)) = -x(t)^\intercal E(t) x(t)$.

2) Minimum exit-norm in dimension 2:

$$T \text{ is free},\ X = U = B_2,\ X_T = \partial X,\ f = u.$$

The Lagrangian to recover is $\bar{L} = \|x\|_2^2 + \|u\|_2^2$. The associated direct OCP is called the minimum exit-norm problem because at $T^\star$, $x(T^\star) \in \partial X$. The optimal control law is $u = x$ and the value function is the polynomial $J^\star_{\bar{L}}(x) = 1 - \|x\|_2^2$.

3) Minimum exit-time in dimension 2:

$$T \text{ is free},\ X = U = B_2,\ X_T = \partial X,\ f = u.$$

The Lagrangian to recover is $\bar{L} = 1$. The optimal control law is $u = x/\|x\|_2$ and the value function is $J^\star_{\bar{L}}(x) = 1 - \|x\|_2$. It corresponds to the problem $(P)_{L_0}$ introduced in Section III-C.

4) Minimum time Brockett integrator:

$$T \text{ is free},\ X = 3B_2,\ U = B_2,\ X_T = \{0\},$$
$$f = (u_1, u_2, x_2 u_1 - x_1 u_2)^\intercal.$$

The Lagrangian to recover is $\bar{L} = 1$. The optimal control law and the value function are described in [19].

C. Numerical results

According to our optimization inverse problem ($IOCP_{a,b}$), we consider several classes of polynomials $\mathcal{L}_{a,b}$ in which we seek a Lagrangian L solution of (IOCP). We vary also the degree of the test function φ. Results of the problems V-B.1, V-B.2, V-B.3, V-B.4 are respectively given in Table I, Table II, Table III and Table IV.
We generate $N = 500$ data points from a set $\mathcal{D} \subset X$ and our algorithm implements the normalization constraint (14).

$\deg \varphi$	class of L	$\varepsilon^\star$	L
4	$\mathcal{L}_{1,1}$	$7e-2$	$0.78x_1^2 + 0.82x_1x_2 + 2.11x_2^2 + 1.12u^2$
10	$\mathcal{L}_{1,0}$	$3.1e-1$	$2.67x_1^2 - 2.31x_1x_2 + 1.33x_2^2$
10	$\mathcal{L}_{1,1}$	$4.5e-6$	$2x_1^2 + 0.5x_1x_2 + x_2^2 + u^2$
10	$\mathcal{L}_{2,2}$	$4.5e-6$	$2x_1^2 + 0.5x_1x_2 + x_2^2 + u^2$

TABLE I

SOLUTION $(L, \varepsilon^\star)$ OF THE PROBLEM $(IOCP)$ ASSOCIATED WITH THE PROBLEM V-B.1.

$\deg \varphi$	class of L	$\varepsilon^\star$	L
2	$\mathcal{L}_{1,1}$	0	$x_1^2 + x_2^2 + u_1^2 + u_2^2$
2	$\mathcal{L}_{2,2}$	0	$x_1^2 + x_2^2 + u_1^2 + u_2^2$
4	$\mathcal{L}_{1,1}$	0	$x_1^2 + x_2^2 + u_1^2 + u_2^2$
4	$\mathcal{L}_{2,2}$	0	$x_1^2 + x_2^2 + u_1^2 + u_2^2$
2	$\mathcal{L}_{0,1}$	$2e-3$	$1.97 + 0.54(u_1^2 + u_2^2)$

TABLE II

SOLUTION $(L, \varepsilon^\star)$ OF THE PROBLEM $(IOCP)$ ASSOCIATED WITH THE PROBLEM V-B.2.

$S_{\mathcal{D}}$	$\deg \varphi$	class of L	$\varepsilon^\star$	L
B_2	4	$\mathcal{L}_{0,1}$	$1e-1$	$0.31 + 0.34u_1^2 + 0.36u_2^2$
B_2	12	$\mathcal{L}_{0,1}$	$2e-2$	$0.327 + 0.335u_1^2 + 0.337u_2^2$
$B_2 \backslash 1/2 B_2$	12	$\mathcal{L}_{0,1}$	$2e-4$	$0.338 + 0.326u_1^2 + 0.336u_2^2$
B_2	2	$\mathcal{L}_{1,1}$	$4.5e-2$	$0.337u_1^2 + 0.339u_2^2$ $+0.741x_1^2 + 0.738x_2^2$
B_2	12	$\mathcal{L}_{1,1}$	$3e-4$	$x_1^2 + x_2^2$
B_2	4	$\mathcal{L}_{0,2}$	0	$(1 - u_1^2 - u_2^2)^2$
B_2	4	$\mathcal{L}_{2,2}$	0	$(1 - u_1^2 - u_2^2)^2$

TABLE III

SOLUTION $(L, \varepsilon^\star)$ OF THE PROBLEM $(IOCP)$ ASSOCIATED WITH THE PROBLEM V-B.3. $S_{\mathcal{D}} \subset X$ IS THE SAMPLE FROM WHICH $\mathcal{D}$ IS GENERATED.

D. Discussion

Clearly, two Lagrangians are equivalent up to a multiplicative constant. This multiplicative constant is set by the value of the constant C in (14).

$\deg\varphi$	class of L	$\varepsilon^\star$	L
10	$\mathcal{L}_{0,1}$	$8.31e-2$	$0.313+0.339\,u_1^2+0.348\,u_2^2$
14	$\mathcal{L}_{0,1}$	$4.36e-2$	$0.323+0.338\,u_1^2+0.339\,u_2^2$
10	$\mathcal{L}_{0,2}$	0	$(1-u_1^2-u_2^2)^2$
10	$\mathcal{L}_{2,2}$	0	$(1-u_1^2-u_2^2)^2$
12	$\mathcal{L}_{1,1}$	$1e-1$	$m_1(x)^\intercal C_1^x m_1(x)+0.31u_1^2+$ $0.35+0.33u_2^2$, $\|C_1^x\|=O(1e-2)$

TABLE IV

SOLUTION $(L,\varepsilon^\star)$ OF THE PROBLEM $(IOCP)$ ASSOCIATED WITH THE PROBLEM V-B.4.

- Problem V-B.1: Table I gives the couple $(\varepsilon^\star, L)$ part of the solution of the inverse problem (IOCP) for several values of the degree of the polynomial φ and several classes of polynomials for the Lagrangian L. The value function $J^\star_{\bar L}$ associated with the Lagrangian $\bar L$ given in Problem V-B.1 is quadratic in x, but it depends also on the time t and we don't know, *a priori*, if the dependance with respect to t of $J^\star_{\bar L}(t,x)$ is polynomial. This provides a reason why L is not a good approximation of $\bar L$ when φ is of degree 4 in t,x. Contrary to the method presented in [17], increasing the degree of φ up to 10, we are able to get exactly $L=\bar L$ with $\varepsilon^\star=0$, provided that the monomial basis for L contains the monomials $u^2, x_1x_2, x_1^2, x_2^2$. If we remove some of these monomials, $\varepsilon^\star$ is bigger. Thus it provides a way to test whether the class $\mathcal{L}_{a,b}$ chosen for L is relevant. The smaller is the optimal value $\varepsilon^\star$ the more relevant is $\mathcal{L}_{a,b}$.
- Problem V-B.2: The Lagrangian $\bar L(x,u)$ is polynomial in x,u and the value function $J_{\bar L}(x)$ is a polynomial of degree 2 in x. Numerically, from Table II taking φ as a polynomial of degree 2 in x, the Lagrangian L solution of (IOCP) corresponds exactly to $\bar L$ provided that the basis of L contains the monomials u_i^2, x_i^2, $i=1,2$. As soon as one of these monomials is removed, the error $\varepsilon^\star$ increases, indicating that the (smaller) class $\mathcal{L}_{a,b}$ for the Lagrangian L is not relevant.
- Problem V-B.3: The control associated with an optimal trajectory satisfies the algebraic equation $u_1^2+u_2^2=1$. The Lagrangian $L=a\,(1+u_1^2+u_2^2)$, $a\neq 0$ is equivalent to $\bar L=1$. From Table III, if the basis of the Lagrangian L contains the monomials $1, u_i, u_i^2$, $i=1,2$, we recover the minimum time problem V-B.3 with a reasonable value of $\varepsilon^\star$ which decreases if the degree of φ increases. Note also that due to the singularity of the value function $J^\star_{\bar L}=1-\|x\|_2$ at $x=0$, $J^\star_{\bar L}$ is hard to approximate by a polynomial near $(0,0)$, hence the objective value $\varepsilon^\star$ of (IOCP) with sample from $S_{\mathcal{D}}=B_2\setminus(1/2\,B_2)$ is smaller.
  Due to the fact that the optimal control $u=(u_1,u_2)$ satisfies $1-u_1^2-u_2^2=0$, a natural Lagrangian is $L=(1-u_1^2-u_2^2)^2$ associated with the zero value function and obtained by our numerical procedure when the dictionary $\mathcal{L}_{a,b}$ contains the monomials $1, u_i^2, u_i^4$, $i=1,2$. This "natural" Lagrangian is purely mathematical with no physical meaning.
  Note that Problem V-B.3 corresponds to $(P)_{L_0}$, $L_0=1=\bar L$ which is defined in Section III-C and has the same optimal control law than problems $(P)_{L_p}$, $p\in\mathbb{N}$. Consider the problems $(P)_{L_1}$ and $(P)_{L_2}$. Note that, in these two problems, one function among L_p and $J^\star_{L_p}$, $p=1,2$ is a polynomial while the other is not differentiable at 0.
  - $(P)_{L_1}$: The Lagrangian to recover is $L_1(x)=\|x\|_2$ associated with the value function $J^\star_{L_1}(x)=1/2\,(1-\|x\|_2^2)$.
    From Table III, in the case $\deg\varphi=2$ and $L\in\mathcal{L}_{1,1}$, the numerical solution (φ,L) of the (IOCP) is $\varphi(x)=1/2\,(1-x_1^2-x_2^2)$ and $L(x,u)=0.337u_1^2+0.339u_2^2+0.741x_1^2+0.738x_2^2$ that we identify to $L(x,u)=0.34+0.74\,(x_1^2+x_2^2)$ since $u_1^2+u_2^2=1$. L corresponds to an approximation on $X=B_2$ of the Lagrangian L_1. Indeed, using MATLAB's routine *fminunc*, a numerical solution of
    $$\min_{a\in\mathbb{R}^6}\int_X \mid a_1+a_2x_1+a_3x_2+a_4x_1^2+a_5x_1x_2+a_6x_2^2 - \sqrt{x_1^2+x_2^2}\mid \mathrm{d}x_1\mathrm{d}x_2$$
    is $a^\star=(0.317,0,0,0.7321,0,0.7321)$ whose values are close to the coefficients of L.
  - $(P)_{L_2}$: The Lagrangian to recover is $L_2(x)=\|x\|_2^2$ associated with the value function $J^\star_{L_2}(x)=1/3\,(1-\|x\|_2^3)$.
    From Table III, in the case $\deg\varphi=12$ and $L\in\mathcal{L}_{1,1}$, the numerical solution L of the (IOCP) is $L(x,u)=x_1^2+x_2^2$. In this case, L corresponds to the Lagrangian L_2 associated with the value function $J^\star_{L_2}$.
- Problem V-B.4: The optimal control satisfies $1-u_1^2-u_2^2=0$. From Table IV, if the basis contains of $\mathcal{L}_{a,b}$ the monomials $1, u_i^2$, $i=1,2$, we recover the Lagrangian of the minimum time since the Lagrangian $L=1+u_1^2+u_2^2$ is equivalent to the Lagrangian $2\bar L$. If we remove some monomials among $1, u_i^2$, $i=1,2$, the value of $\varepsilon^\star$ increases, which again invalidates the choice of the smaller class $\mathcal{L}_{a,b}$ for L.
  If the basis contains the monomials $1, u_i^2, u_i^4$, $i=1,2$ we recover the Lagrangian $L=(1-u_1^2-u_2^2)^2$ associated with the zero value function (see Section III-C).

It is important to point out that we are able to recover the Lagrangian $\bar L$ without adding a regularization parameter which was necessary in [18]. Indeed in [18] the authors adopted a more general point of view where $\mathcal{L}=\mathbb{R}[x,u]_d$ with d relatively large. They minimize a trade-off between the error $\varepsilon^\star$ and the value of a regularization parameter controlling the *sparsity* of the polynomial L: the more one asks for sparsity, the larger is the resulting optimal error $\varepsilon^\star$ and vice versa. From an applicative point of view, the main drawback of this method is that

we do not know an *a priori* value of the regularization parameter for which the Lagrangian $\bar{L}$ is recovered. In this sense, the method described in this paper – where a particular a class of Lagrangian functions is imposed – seems more suitable for applications.

VI. Conclusion

Hamilton-Jacobi-Bellman conditions provide a global-optimality certificate which is a powerful and ideal tool for solving inverse optimal control problems. While intractable in full generality, it can be implemented when data of the control system are polynomials and semi-algebraic sets, in the spirit of [18]. Indeed powerful positivity certificates from real algebraic geometry allow to implement a relaxed version of HJB which imposes (i) positivity constraints on the unknown value function and Lagrangian, and (ii) a guaranteed ε-global optimality for all the given trajectories of the database $\mathcal{D}$. In doing so we can solve efficiently inverse optimal control problems of relatively small dimension.

Compared to [17], our method is less general as the search of an optimal Lagrangian is done on some (restricted) class of Lagrangians defined *a priori*. On the other hand, the search is more efficient with *no* need of a regularizing parameter to control the sparsity. When considering the same examples as in [18], we obtain more accurate estimations on Lagrangians without a sparsity constraint on the Lagrangian. Finally, as an additional and interesting feature of the method, in a way a small resulting optimal value $\varepsilon^\star$ validates the choice of the class of Lagrangians considered, while a relatively large $\varepsilon^\star$ is a strong indication that the class is too small.

Future investigations will try to determine if our numerical scheme provides interesting results on practical inverse problems, especially those coming from humanoid robotics, an important field of application [15] for inverse optimal control.

VII. ACKNOWLEDGMENTS

This work was supported by the European Research Council (ERC) through an ERC-Advanced Grant for the TAMING project. The authors would like to thank Edouard Pauwels for helpful discussions based on upstream works.

References

[1] A.A. Ahmadi, A. Majumdar. DSOS and SDSOS Optimization: LP and SOCP-Based Alternatives to SOS Optimization, Proceedings of the 48th Annual Conference on Information Sciences and Systems, Princeton, NJ, 1–5, March 2014.

[2] A. Ajami, J.-P. Gauthier, T. Maillot, U. Serres. How humans fly. ESAIM: Control, Optimisation and Calculus of Variations, 19(04):1030–1054, 2013.

[3] M. Bardi, I. Capuzzo-Dolcetta. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Springer, 2008.

[4] D.P. Bertsekas. Dynamic programming and optimal control Vol. II. Athena Scientific, Belmont, MA, 2001.

[5] F.C. Chittaro, F. Jean, P. Mason. On inverse optimal control problems of human locomotion: Stability and robustness of the minimizers. Journal of Mathematical Sciences, 195(3):269–287, 2013

[6] F. Jean. Control of Nonholonomic Systems: from Sub-Riemannian Geometry to Motion Planning. Springer International Publishing, SpringerBriefs in Mathematics, 2014

[7] F. Jean, S. Maslovskaya, I. Zelenko. Inverse Optimal Control Problem: the Sub-Riemannian Case. (preprint http://hal.inria.fr/hal-01452458/document)

[8] R. E. Kalman. The Theory of Optimal Control and the Calculus of Variations, in R. Bellman (ed.), "Mathematical Optimization Techniques," University of California Press, Berkeley, Calif., 1963.

[9] J. B. Lasserre. Global optimization with polynomials and the problem of moments. SIAM Journal on Optimization, 11(3):796–817, 2001.

[10] J. B. Lasserre. Moments, positive polynomials and their applications. Imperial College Press, UK, 2010.

[11] J. B. Lasserre, Kim-Chuan Toh, Shouguang Yang. A bounded degree SOS hierarchy for polynomial optimization" European J. Comput. Optim. 5, 87–117, 2017.

[12] J.-P. Laumond. Robot Motion Planning and Control, Springer Verlag, 1998.

[13] J.-P. Laumond, N. Mansard and J.B. Lasserre. Geometric and Numerical Foundations of Movements, Springer Tracts in Advanced Robotics, New York, 2017

[14] J. Löfberg. Pre-and post-processing sum-of-squares programs in practice. IEEE Transactions on Automatic Control, 54(5):1007–1011, 2009.

[15] K. Mombaur, A. Truong, J. P. Laumond. From human to humanoid locomotion–an inverse optimal control approach. Autonomous robots, 28(3):369–383, 2010.

[16] MOSEK ApS. The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 63)", year = 2016, url = "http://docs.mosek.com/7.1/toolbox/index.html"

[17] E. Pauwels, D. Henrion, J.-B. Lasserre. Inverse optimal control with polynomial optimization. IEEE Conference on Decision and Control, Los Angeles, December 2014.

[18] E. Pauwels, D. Henrion, J.-B. Lasserre. Linear conic optimization for inverse optimal control. SIAM J. Control & Optim. 54:1798–1825, 2016.

[19] C. Prieur, E. Trélat. Robust optimal stabilization of the Brockett integrator via a hybrid feedback. Mathematics of Control, Signals and Systems, 17(3), 201–216, 2005.

[20] M. Putinar. Positive polynomials on compact semi-algebraic sets. Indiana University Mathematics Journal, 42(3):969–984, 1993.

[21] R. Vinter, Optimal control. Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA xviii+507, 2000