Linear conic optimization for inverse optimal control

Jean-Bernard Lasserre

Linear conic optimization for inverse optimal control

Edouard Pauwels[1], Didier Henrion[2,3,4], Jean-Bernard Lasserre[2,3]

Draft of November 5, 2014

Abstract

We address the inverse problem of Lagrangian identification based on trajectories in the context of nonlinear optimal control. We propose a general formulation of the inverse problem based on occupation measures and complementarity in linear programming. The use of occupation measures in this context offers several advantages from the theoretical, numerical and statistical points of view. We propose an approximation procedure for which strong theoretical guarantees are available. Finally, the relevance of the method is illustrated on academic examples.

1 Introduction

In the context of nonlinear optimal control, we are interested in the inverse problem of Lagrangian identification from given trajectories. This identification should be carried out such that solving the direct optimal control problem with the identified Lagrangian would allow to recover the given trajectories.

Inverse problems of calculus of variations are old topics that have attracted a renewal of interest in the context of optimal control, especially in humanoid robotics [4]. Relevant aspects of the problem are not well understood and many issues still need to be addressed to propose a tool that could be used in experimental settings. The work presented here constitutes a step in this direction. A preliminary conference version [27] originally introduced our optimization framework as a tool to solve the inverse problem numerically. The current paper extends this work in many ways. In particular, by using the (quite general) concept of *occupation measures* we can propose a broad definition of inverse optimality and we also rigorously justify most of the approximations behind the numerical results reported in [27]. Many aspects of this work parallel the results of [21] about *direct* optimal control with polynomial data.

[1]Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa 3200, Israel.

[2]CNRS; LAAS; 7 avenue du colonel Roche, F-31400 Toulouse; France.

[3]Université de Toulouse; LAAS, F-31400 Toulouse, France.

[4]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czech Republic

1.1 Context

Since its introduction by Kalman [18], the inverse problem of optimal control has been studied in linear settings [3, 17, 14, 26] leading to many nonlinear variations [32, 25, 10, 13]. In these works the input data of the problem is a characteristic of a class of trajectories often given in the form of a control law. This contrasts with the setting we propose to study, for which the input is a set of trajectories which could come from physical experiments. This motivates the work of [11] and [2] about well-posedness of the inverse problem, both in the context of unicycle dynamics in robotics and strictly convex positive Lagrangians.

On the other hand, to treat the inverse problem several authors have proposed numerical methods based on the ability to solve the direct problem [24], also in the context of Markov decision process [1, 31] or based on a discretized version of the direct problem [30, 19].

Our approach is different and based on occupation measures, an abstract and quite general tool to handle trajectories (or limiting trajectories) of feasible solutions of classical control problems. Formulating the (direct) control problem on appropriate spaces of measures amounts to relaxing the original problem. In most applications, both relaxed and original problems have same optimal value [36, 35, 15]. However the relaxed formulation has the crucial advantage that compactness holds in a certain weak sense: As a matter of fact, many optimization problems over appropriate spaces of measures attain their optimum, whereas most optimization problems over smaller functional spaces (e.g. continuous functions, or Lebesgue integrable functions) typically have *no* optimal solution. At last but not least, for control problems with polynomial data, the relaxed problem can be formulated as an optimization problem on *moments* of occupation measures. By combining this with relatively recent advances in real algebraic geometry [29] and in numerical optimization [20] one may thus provide a systematic numerical scheme to approximate effectively relaxed solutions of optimal control problems [21].

1.2 Contribution

We choose the setting of free terminal time optimal control which is consistent with many physical experiments that one can think of. But the same approach with *ad hoc* modifications is also valid in the fixed terminal time setting.

• In our opinion, occupation measures are the perfect abstract tool to formally express the fact that we consider a (possibly uncountably infinite) superposition of trajectories as input data of the inverse control problem. We then propose a general formulation of the inverse problem based on occupation measures and complementarity in linear programming. A relaxation of the well known Hamilton-Jacobi-Bellman (HJB) sufficient optimality condition appears in our formulation as for the usual direct optimal control problem [16]. This formulation is shown to be consistent with what is commonly expected regarding inverse optimality.

It is worth noting that when using the HJB optimality conditions, the situation is completely symmetric for the direct and inverse control problems. In both cases the HJB optimality conditions are used to *certify the global optimality of trajectories*. But in the

former the Lagrangian is known and HJB provide conditions on the optimal state-control trajectories (to be determined) whereas in the latter the "optimal" state-control trajectories are known and HJB provide conditions on the Lagrangian (to be determined) for the given trajectories to be optimal. (In both cases the optimal value function is considered as an auxiliary "variable".)

• Furthermore, this framework allows to further characterize the space of solutions associated with a given inverse optimal control problem. This viewpoint is different from what has been proposed in previous (theoretical and numerical) contributions to this problem [24, 30, 11, 2] which, implicitly or explicitly, involve strong (and, in our opinion, overly restrictive) constraints on the class of functions in which the candidate Lagrangians are searched.

• Remarkably, despite the abstract setting of occupation measures, the proposed formulation is amenable to explicit numerical approximations via a hierarchy of semi-definite programs[1]. Indeed in the context of polynomial dynamics and semi-algebraic constraints, both the optimal value function and Lagrangian used in the (relaxed) HJB optimality conditions can be approximated with polynomials. We show that such a reinforcement is coherent in the sense that no polynomial solution to the inverse problem is lost.

• Finally, in usual experimental settings one does not have access to complete trajectories. Instead one is rather given finitely many data points sampled from trajectories. But results from probability applied to our occupation measures allow to formalize the fact that we only work with "samples". In addition, in this framework one may use empirical processes and statistical learning theory [34, 9] to provide bounds on the error made when working with samples instead of original trajectories.

Organization of the paper. In Section 2 we provide the context and background on optimal control and occupation measures. In Section 3, we present our characterization of solutions to the inverse optimal control problem and illustrate how it allows to further discuss about the set of solutions and links with the direct optimal control problem. Numerical approximations via polynomials and statistical approximations via finite samples are provided and discussed in Section 4. The resulting numerical scheme (with proven strong theoretical guarantees) can be implemented with off-the-shelf software on a standard computer. Finally, Section 5 describes numerical results on academic examples.

2 Preliminaries

2.1 Notations

If A is a compact subset of a finite-dimensional Euclidean space, let $\mathcal{C}(A)$ resp. $\mathcal{C}^1(A)$ denote the set of continuous resp. continuously differentiable functions from A to $\mathbb{R}$. Let $\mathcal{M}(A)$ denote the space of Borel measures on A, the topological dual of $\mathcal{C}(A)$ with duality

[1] A semi-definite program is a finite-dimensional linear optimization problem over the cone of non-negative quadratic forms for which powerful primal-dual interior-point algorithms are available [33].

bracket denoted by $\langle ., . \rangle$, i.e. $\langle \mu, f \rangle = \int_A f(x) d\mu(x)$ is the integration on A of a function $f \in \mathcal{C}(A)$ with respect to a measure $\mu \in \mathcal{M}(A)$. Let $\mathcal{M}_+(A)$ resp. $\mathcal{C}_+(A)$ denote the cone of non-negative Borel measures resp. non-negative continuous functions on A. The support of a measure $\mu \in \mathcal{M}_+(A)$ is denoted by $\operatorname{spt} \mu$. An element $\mu \in \mathcal{M}_+(A)$ such that $\langle \mu, 1 \rangle = 1$ is called a probability measure. Let δ_x denote the Dirac measure concentrated on x and let $I(e)$ denote the indicator function of an event e, equal to 1 if e is true, and 0 otherwise.

Let $X \subseteq \mathbb{R}^{d_X}$ denote the state space and $U \subseteq \mathbb{R}^{d_U}$ denote the control space which are supposed to be compact subsets of Euclidean spaces. System dynamics are given by a continuously differentiable vector field $f \in \mathcal{C}^1(X \times U)^{d_X}$. Terminal state constraints are modeled by a set $X_T \subset X$ which is also given. Let B_n denote the unit ball of the Euclidean norm in $\mathbb{R}^n$, and let ∂S denote the boundary of set S in the Euclidean space. Let $\mathbb{R}[z]$ denote the set of multivariate polynomials with real coefficients with variables z and let $\mathbb{R}_k[z]$ denote the set of such polynomials with degree at most k. For a polynomial $p \in \mathbb{R}_k[z]$, we denote by $\|p\|_1$ the sum of the absolute values of the coefficients of p when expanded in the monomial basis.

2.2 Context: free terminal time optimal control

We consider direct optimal control problems of the form:

$$\begin{aligned} v_0(z) := \; & \inf_{u,T} \int_0^T l_0(x(t), u(t)) dt \\ & \text{s.t. } \dot{x}(t) = f(x(t), u(t)), \\ & \quad x(t) \in X,\, u(t) \in U,\, t \in [0, T], \\ & \quad x(0) = z,\, x(T) \in X_T, \\ & \quad T \in [0, T_M] \end{aligned} \tag{ocp$_0$}$$

with Lagrangian $l_0 \in \mathcal{C}(X \times U)$ and free final time T with a given upper bound T_M which ensures that the value function v_0 is bounded below. Dynamics f are given, as well as the sets X, U and $X_T \subset X$. We assume that a set $X_0 \subset X$ is given such that the following assumption is satisfied:

Assumption 1 *For all initial conditions $z \in X_0$, problem (ocp$_0$) is feasible.*

2.3 Occupation measures

The *occupation measure* and *terminal measure* of a set of trajectories are defined as follows. Consider a probability measure $\mu_0 \in \mathcal{M}_+(X_0)$. Suppose that for each $z \in \operatorname{spt} \mu_0$, there exists a terminal time $T_z \in [0, T_M]$, a measurable control $u_z \colon [0, T_z] \to U$ and an absolutely continuous trajectory $x_z \colon [0, T_z] \to X$ such that

$$\begin{aligned} \dot{x}_z(t) &= f(x_z(t), u_z(t)), \\ x_z(0) &= z, \;\; x_z(T_z) \in X_T. \end{aligned} \tag{1}$$

The occupation measure $\mu \in \mathcal{M}_+(X \times U)$ and terminal measure $\mu_T \in \mathcal{M}_+(X_T)$ of the set of trajectories $\{x_z(t)\}_{z \in \text{spt}\,\mu_0, t \in [0,T_z]}$ are then defined as follows:

$$\begin{aligned} \mu(A \times B) &:= \int_{X_0} \left(\int_0^{T_z} I(x_z(t) \in A) I(u_z(t) \in B) dt \right) \mu_0(dz), \\ \mu_T(A) &:= \int_{X_0} I(x_z(T_z) \in A)\, \mu_0(dz), \end{aligned} \tag{2}$$

for every Borel sets $A \subset X$ and $B \subset U$. With the previous definition,

$$\langle \mu, l \rangle = \int_{X_0} \left(\int_0^{T_z} l(x_z(t), u_z(t)) dt \right) \mu_0(dz), \quad \forall l \in \mathcal{C}(X \times U).$$

In particular

$$\mu(X \times U) = \langle \mu, 1 \rangle = \int_{X_0} T_z\, \mu_0(dz).$$

Furthermore for every $v \in \mathcal{C}^1(X)$,

$$\begin{aligned} \langle \mu, \text{grad}\, v \cdot f \rangle &= \int_{X_0} \left(\int_0^{T_z} \text{grad}\, v(x_z(t)) \cdot f(x_z(t), u_z(t))\, dt \right) \mu_0(dz) \\ &= \int_{X_0} (v(x_z(T_z)) - v(x_z(0)))\, \mu_0(dz) \\ &= \langle \mu_T, v \rangle - \langle \mu_0, v \rangle, \end{aligned} \tag{3}$$

where "grad" denotes the gradient vector of first order derivatives of v, and the "dot" denotes the inner product between vectors. Equation (3) is known as Liouville's equation and is also written as

$$\text{div} f\mu + \mu_T = \mu_0, \tag{4}$$

where the divergence is to be interpreted in the weak sense and a change of sign comes from integration by part. As we have seen, occupation and terminal measures as defined in (2) satisfy the Liouville equation (4). The converse is not true, not all occupation measures that satisfy the Liouville equation (4) arise in this way. However, for inverse optimal control, we suppose that the occupation measure is given. Moreover, the Liouville equation and positivity constraints are sufficient to develop some theoretic aspects of inverse optimality.

Therefore, independently of how it is constructed, the input data of our inverse control problem is an occupation measure satisfying the Liouville equation.

All the results will in particular apply to situations when the occupation measure is constructed as described in this section.

2.4 Direct optimal control

Using the formalism of occupation measures, given a continuous Lagrangian l, an initial measure μ_0 and a maximal terminal time T_M, we consider direct optimal control problems

of the form

$$\begin{aligned} p^*(\mu_0) := \quad & \inf_{\mu,\mu_T} \ \langle \mu, l \rangle \\ & \text{s.t.} \ \ \mathrm{div} f\mu + \mu_T = \mu_0, \\ & \quad \langle \mu, 1 \rangle \leq T_M, \\ & \quad \mu \in \mathcal{M}_+(X \times U), \\ & \quad \mu_T \in \mathcal{M}_+(X_T). \end{aligned} \tag{ocp}$$

Definition 1 (OCP) $\mathrm{OCP}(l, \mu_0, T_M)$ *is the set of measures* (μ, μ_T) *solving problem (ocp).*

Note that by Lemma 1 and Assumption 1, set $\mathrm{OCP}(l, \mu_0, T)$ is not empty. The link between problems (ocp_0) and (ocp) is far from trivial. The decision variables (μ, μ_T) in (ocp) can be viewed as a superposition of limiting objects arising from problem (ocp_0). It is possible to construct problems for which measures considered in problem (ocp) do not arise in this way and this may introduce spurious minimizers which are far from classical trajectories of problem (ocp_0). These problems are usually overly constrained and not physically relevant, and in most practical settings, we have

$$p^*(\delta_z) = v_0(z) \qquad \forall\, z \in \mathrm{spt}\, \mu_0,$$

which we could see as an assumption on the inverse problem data. A sufficient condition is given in [36, Theorem 2.3]

All our analysis refers to direct control problems of the form of (ocp). Its conic dual can be written as

$$\begin{aligned} d^*(\mu_0) := \quad & \sup_{v,w} \ \langle \mu_0, v \rangle - wT_M \\ & \text{s.t.} \ \ l + w + \mathrm{grad}\, v \cdot f \in \mathcal{C}_+(X \times U) \\ & \quad -v \in \mathcal{C}_+(X_T), \\ & \quad w \geq 0, \\ & \quad v \in \mathcal{C}^1(X), \ \ w \in \mathbb{R}. \end{aligned} \tag{hjb}$$

The first two constraints $l + w + \mathrm{grad}\, v \cdot f \in \mathcal{C}_+(X \times U)$ and $-v \in \mathcal{C}_+(X_T)$ of (hjb) are relaxations of the well-known Hamilton-Jacobi-Bellman (HJB) sufficient condition of optimality [5, 6]. Conic duality provides the following link between the problems (ocp) and (hjb).

Lemma 1 *The infimum in (ocp) is attained and there exists a maximizing sequence in (hjb). In addition, for any feasible primal pair* (μ, μ_T) *and any sequence of dual variables* $v_k \in \mathcal{C}^1(X)$ *and* $w_k \in \mathbb{R}$, $k \in \mathbb{N}$, *the following assertions are equivalent*

- (μ, μ_T) *is optimal for (ocp) and* $(v_k, w_k)_{k\in\mathbb{N}}$ *is a maximizing sequence for (hjb);*
- *strong duality:*

$$\langle \mu_0, v_k \rangle - w_k T_M \underset{k\to\infty}{\longrightarrow} \langle \mu, l \rangle\,; \tag{5}$$

- *complementarity:*

$$\begin{aligned} w_k(\langle \mu, 1 \rangle - T_M) &\underset{k\to\infty}{\longrightarrow} 0, \\ \langle \mu, l + w_k + \mathrm{grad}\, v_k \cdot f \rangle &\underset{k\to\infty}{\longrightarrow} 0, \\ \langle \mu_T, v_k \rangle &\underset{k\to\infty}{\longrightarrow} 0. \end{aligned} \tag{6}$$

Proof : We only sketch the proof here, for more details see [21]. Observe that (ocp) is feasible thanks to Assumption 1 and (hjb) is feasible with $w = \max(-\min_{X\times U} l, 0)$ and $v = 0$. Moreover, the cone $\{(\mathrm{div} f\mu + \mu_T, \langle \mu, 1\rangle, \langle \mu, l_0\rangle) : \mu \in \mathcal{M}_+(X\times U), \mu_T \in \mathcal{M}_+(X_T)\}$ is closed for the weak topology $\sigma(\mathcal{M}_+(X)\times R^2, C_+(X)\times R^2)$ (by using Banach-Alaoglu's Theorem). Therefore there is no duality gap between (ocp) and (hjb) and the optimum is attained in the primal, see e.g. [7, Theorem IV.7.2]. Condition (5) is just a reformulation of strong duality in this context. Equivalence with (6) follows by noticing that for any primal feasible pair (μ, μ_T) and dual feasible pair (v, w),

$$\begin{aligned}
\langle \mu, l\rangle &= \langle \mu, l\rangle - \langle \mathrm{div} f\mu + \mu_T - \mu_0, v\rangle \\
&\geq \langle \mu, l\rangle - \langle \mathrm{div} f\mu + \mu_T - \mu_0, v\rangle + w(\langle \mu, 1\rangle - T_M) \\
&= \langle \mu, l + w + \mathrm{grad}\, v \cdot f\rangle - \langle \mu_T, v\rangle + \langle \mu_0, v\rangle - wT_M \\
&\geq \langle \mu_0, v\rangle - wT_M.
\end{aligned}$$

□

Remark 1 *If the Lagrangian l is strictly positive on $X\times U$, then Lemma 1 holds without the constraint $\langle \mu, 1\rangle \leq T_M$ and without the dual variable w.*

3 Inverse optimal control

Given a "set" of trajectories and model constraints, the inverse problem of optimal control consists of finding a Lagrangian for which the trajectories are optimal. Thanks to the framework exposed in the previous section, it is now easy to define what is a solution to the inverse optimal control problem.

• Firstly, the "set" of trajectories will be represented by measures satisfying Liouville equation (4) which are part of the data of the inverse problem.

• Secondly, a Lagrangian l solution to the inverse problem is a continuous function such that $(\mu, \mu_T) \in \mathrm{OCP}(l, \mu_0, T)$ for some T such that $\mathrm{OCP}(l, \mu_0, T)$ is feasible.

In this section, we propose a rigorous definition of inverse optimality and prove an equivalence result between direct and inverse optimality. To do so, we use Lemma 1 which ensures that $\mathrm{OCP}(l, \mu_0, T)$ is non empty as long as $T \leq T_M$. Furthermore, it provides a certificate of (sub)optimality.

3.1 What is a solution to the inverse optimal control problem?

We can now formally define what is meant by a solution to the inverse optimal control problem:

Definition 2 (IOCP and IOCP$_\epsilon$) *For $\epsilon > 0$, given measures $\mu \in \mathcal{M}_+(C\times U)$ and $\mu_T \in \mathcal{M}_+(X_T)$ such that $\mathrm{div} f\mu + \mu_T \in \mathcal{M}_+(X_0)$, denote by $\mathrm{IOCP}_\epsilon(\mu, \mu_T)$ the set of ϵ-optimal solutions to the inverse optimal control problem, namely the set of functions*

$l \in \mathcal{C}(X \times U)$ *such that there exists a function* $v \in \mathcal{C}^1(X)$ *satisfying*

$$\begin{array}{rcl} \langle \mu, l + \operatorname{grad} v \cdot f \rangle & \leq & \epsilon, \\ l + \operatorname{grad} v \cdot f + \epsilon & \in & \mathcal{C}_+(X \times U), \\ \langle \mu_T, v \rangle & \geq & -\epsilon, \\ -v & \in & \mathcal{C}_+(X_T). \end{array}$$

Then the set $\mathrm{IOCP}(\mu, \mu_T)$ *of solutions to the inverse optimal control is defined by:*

$$\mathrm{IOCP}(\mu, \mu_T) := \{ l \in \mathcal{C}(X \times U) : \ l \in \mathrm{IOCP}_\epsilon(\mu, \mu_T) \quad \forall \epsilon > 0 \}.$$

The motivation behind the above definition is the following:

Theorem 1 *Given* $\mu \in \mathcal{M}_+(C \times U)$ *and* $\mu_T \in \mathcal{M}_+(X_T)$*, the set* $\mathrm{IOCP}(\mu, \mu_T)$ *is a convex cone, closed for the supremum norm. Moreover, the following two assertions are equivalent:*

- $l \in \mathrm{IOCP}(\mu, \mu_T)$, $\operatorname{div} f\mu + \mu_T = \mu_0 \in \mathcal{M}_+(X_0)$*;*
- $\exists T > \langle \mu, 1 \rangle$, $(\mu, \mu_T) \in \mathrm{OCP}(l, \mu_0, T)$.

Proof : Convexity follows from convexity of the constraints of Definition 2. There exists a constant K such that for any pair (l_0, v_0) that satisfies constraints of Definition 2 for a certain $\epsilon > 0$, then it holds for any Lagrangian l that $\langle \mu, l + \operatorname{grad} v_0 \cdot f \rangle \leq \epsilon + K\|l - l_0\|_\infty$ and $l + \operatorname{grad} v_0 \cdot f \geq -\epsilon - K\|l - l_0\|_\infty$ on $X \times U$, which is sufficient to prove closedness.

For the first implication, suppose that $l \in \mathrm{IOCP}(\mu, \mu_T)$ and $\operatorname{div} f\mu + \mu_T = \mu_0$. Then for any $T > \langle \mu, 1 \rangle$, the pair (μ, μ_T) is feasible for $\mathrm{OCP}(l, \mu_0, T)$. Lemma 1 holds and the definition of $\mathrm{IOCP}(\mu, \mu_T)$ allows to construct a dual sequence that is feasible for $\mathrm{OCP}(l, \mu_0, T)$ and that satisfies the complementarity condition with the pair (μ, μ_T).

We now turn to the last implication. Suppose that $(\mu, \mu_T) \in \mathrm{OCP}(l, \mu_0, T)$ and $\langle \mu, 1 \rangle < T$. In particular, (μ, μ_T) is feasible for $\mathrm{OCP}(l, \mu_0, T)$ and Lemma 1 holds. Consider the dual maximizing sequence $(v_k, w_k)_{k \in \mathbb{N}}$ given by Lemma 1. Complementarity ensures that $\lim_{k \to \infty} w_k = 0$. Furthermore, it holds that $\lim_{k \to \infty} \langle \mu, l + \operatorname{grad} v_k \cdot f \rangle = 0$, $l + \operatorname{grad} v_k \cdot f \geq -w_k$ on $X \times U$, $\lim_{k \to \infty} \langle \mu_T, v_k \rangle = 0$ and $v_k \leq 0$ on X_T which shows that $l \in \mathrm{IOCP}(\mu, \mu_T)$. $\square$

Remark 2 Another motivation behind Definition 2 of inverse optimality is the following. Suppose that $l \in \mathrm{IOCP}_\epsilon(\mu, \mu_T)$, then for any $T \geq \langle \mu, 1 \rangle$, (μ, μ_T) is close to optimal for the problem $\mathrm{OCP}(l, \mu_0, T)$. Indeed, suppose that $(\tilde{\mu}, \tilde{\mu}_T) \in \mathrm{OCP}(l, \mu_0, T)$. Then there exists v such that

$$\begin{array}{rcl} \langle \mu, l + \operatorname{grad} v \cdot f \rangle & = & \langle \mu, l \rangle + \langle \mu_T - \mu_0, v \rangle \leq \epsilon \\ \langle \tilde{\mu}, l + \operatorname{grad} v \cdot f \rangle & = & \langle \tilde{\mu}, l \rangle + \langle \tilde{\mu}_T - \mu_0, v \rangle \geq -T\epsilon, \end{array}$$

and $-\langle \mu_T, v \rangle \leq \epsilon$ as well as $\langle \tilde{\mu}_T, v \rangle \leq \epsilon$. In addition, $\langle \mu, l \rangle \leq \langle \mu_0, v \rangle + 2\epsilon$ and $\langle \tilde{\mu}, l \rangle \geq \langle \mu_0, v \rangle - T\epsilon$. Therefore $\langle \mu, l \rangle \geq \langle \tilde{\mu}, l \rangle \geq \langle \mu, l \rangle - (T + 2)\epsilon$.

Remark 3 At first sight the introduction of T in Theorem 1 may look artificial whereas in fact it carries important information. The second part in the equivalence states that (μ, μ_T) is a solution to some direct problem and does not saturate one of the constraints. This allows to avoid direct problems for which, for any value of T, any solution would saturate the constraint on the mass of the occupation measure; for example this happens in direct problems with free terminal time tending to infinity. Such problems should be avoided since then an occupation measure with finite mass cannot be optimal. Given a Lagrangian l, there is no guarantee that there exists a triplet (μ, μ_T, T) which satisfies the second point of Theorem 1. However, checking that a Lagrangian l meets our criterion for inverse optimality ensures that this is the case.

Remark 4 If $l \in \mathrm{IOCP}(\mu, \mu_T)$ and (μ, μ_0, μ_T) is a superposition of classical trajectories as defined in Section 2.3, then μ_0-a.a. (almost all) of these trajectories must be optimal for the corresponding direct problem. In particular, $v_0(z)$ given by (ocp_0) is attained for μ_0-a.a. initial conditions z.

3.2 Applications to inverse optimality

We claim that Definition 2 is a powerful tool to analyze inverse optimality in the context of optimal control. To go beyond Theorem 1, we next describe results and comments that stem from Definition 2 of inverse optimality.

3.2.1 How big is the space of solutions to the inverse problem?

Theorem 1 justifies the idea that if trajectories realize the minimum of some optimal control process then the corresponding Lagrangian meets our criterion. This requirement is necessary for any "inverse problem" (and not only for inverse optimal control). However in general there could be many candidate solutions as illustrated in this section. In what follows, we assume that the triplet (μ, μ_0, μ_T) satisfies Liouville's equation (4).

Conserved values. Suppose that there exists a function $g \in \mathcal{C}(X \times U)$ such that $g(x, u) = 0$ for all $(u, x) \in \mathrm{spt}\, \mu$. Then $g^2 \in \mathrm{IOCP}(\mu, \mu_T)$. In practical examples there might be many such conserved values. For instance this is the case when x or u or both remain on a manifold or when there exists a continuous mapping $x \to u(x)$.

Total variations. Consider any function $g \in \mathcal{C}^1(X)$. All Lagrangians of the form $l = \mathrm{grad}\, g \cdot f$ belong to $\mathrm{IOCP}(\mu, \mu_T)$, independently of (μ, μ_0, μ_T).

Convex conic combinations and uniform limits of solutions. As stated in Theorem 1, the set of solutions to the inverse problem is a convex cone, closed for the supremum norm. For example, let $g \in \mathcal{C}^1(X)$ and consider a Lagrangian $l \in \mathcal{C}(X \times U)$. Then $\mathrm{OCP}(l, \mu_0, T) = \mathrm{OCP}(l + \mathrm{grad}\, g \cdot f, \mu_0, T)$ for every $T > 0$ and therefore both Lagrangians l and $l + \mathrm{grad}\, g \cdot f$ are solutions to the inverse optimal control problem.

All the above examples illustrate that many solutions to the inverse problem may exist. Although these solutions are valid from a theoretical point of view, they do not correspond to what is commonly expected from a solution. Indeed, they do not arise from an optimal physical process that would have generated trajectories, but rather from mathematical artifacts.

3.2.2 How does the direct problem affect the space of solutions to the inverse problem?

Intuitively, the more information is contained in (μ, μ_T, μ_0), the smaller is the space of solutions to the inverse problem. We next discuss two factors that impact the size of $\mathrm{IOCP}(\mu, \mu_T)$.

Direct problem constraints. Denote by $\mathcal{K}^1$ (resp. $\mathcal{K}^2$) the feasible set of problem (ocp) and by $\mathrm{IOCP}^1(\mu, \mu_T)$ (resp. $\mathrm{IOCP}^2(\mu, \mu_T)$) the set of solutions to the inverse problem (as described in Definition 2) when the state, control and dynamical constraints are given by (X^1, U^1, f^1) (resp. (X^2, U^2, f^2)).

If $\mathcal{K}^1 \subset \mathcal{K}^2$ then $\mathrm{IOCP}^2(\mu, \mu_T) \subset \mathrm{IOCP}^1(\mu, \mu_T)$. In other words, there is a kind of duality between the space of feasible solutions for the direct problem and the space of solutions to the inverse problem. An extreme instance is when the feasible space of the direct problem is a singleton (f does not depend on the control u), in which case any Lagrangian is a solution to the inverse problem.

Range of the occupation measure. Suppose that $(\mu, \mu_0, \mu_T) = (\mu^1, \mu_0^1, \mu_T^1) + (\mu^2, \mu_0^2, \mu_T^2)$ where $\mathrm{div} f\mu^1 + \mu_T^1 = \mu_0^1$ and $\mathrm{div} f\mu^2 + \mu_T^2 = \mu_0^2$. Then $\mathrm{IOCP}(\mu, \mu_T) \subset \mathrm{IOCP}(\mu^1, \mu_T^1)$. As a consequence, maximizing the support of the initial measure μ_0 reduces the space of solutions to the inverse problem. When the occupation measure μ is a superposition of trajectories as detailed in Section 2.3, the larger is the "space" occupied by trajectories, the smaller is the space of potential solutions to the inverse problem.

3.2.3 A toy example of quantitative well-definedness analysis

To illustrate the proposed framework we consider a simple uni-dimensional example. We emphasize that his example is very simple in the sense that the direct problem is easy. However, inspecting the solution of the inverse problem leads to non trivial behaviors. Let $X = [-1, 1]$ with $X_T = \{0\}$ and let $f(x, u) = u$ with $U = [-1, 1]$. Consider the family of Lagrangians $\mathcal{F} = \{l_\alpha \colon u \mapsto 1 + \frac{\alpha}{2}u^2,\ \alpha \geq 0\}$. Suppose that we are given a triplet (μ, μ_0, μ_T) which consists of a superposition of trajectories as described in Section 2.3. We wish to find a candidate Lagrangian in the family $\mathcal{F}$. Then we have the following alternatives.

1 $\mathcal{F} \cap \mathrm{IOCP}(\mu, \mu_T) = \emptyset$.

2 $\mathcal{F} \cap \mathrm{IOCP}(\mu, \mu_T) = \mathcal{F}$.

3 $\mathcal{F} \cap \mathrm{IOCP}(\mu, \mu_T)$ is a singleton, $\{l_\alpha\}$, $\alpha > 2$.

4 $\mathcal{F} \cap \mathrm{IOCP}(\mu, \mu_T) = \{l_\alpha,\, 0 \leq \alpha \leq 2\}$.

We should comment on case 1 latter. If the support of μ is empty, which means that $\mu_0 = \mu_T = \delta_x$, then we are in case 2. Assume now that the support of μ is non-empty and we are not in case 1. Then there exists $\alpha \geq 0$ such that $l_\alpha \in \mathrm{IOCP}(\mu, \mu_T)$. Consider a sequence of decreasing positive numbers $\epsilon_k \to 0$ and the corresponding certificates functions v_k that allow to verify that $l_\alpha \in \mathrm{IOCP}_{\epsilon_k}(\mu, \mu_T)$. Since $\mu_T = \delta_0$, we may assume (up to an addition) that $v_k(0) = 0$ which simplifies the problem. In addition, one must have $1 + \frac{\alpha}{2}u^2 + v_k'(x)u \to 0$, μ almost every where (recall that the support of μ is non empty and this concerns a non empty subset of X and U). Furthermore, for any x, one must have $1 + \frac{\alpha}{2}u^2 + v_k'(x)u + \epsilon_k = \left(\sqrt{\frac{\alpha}{2}}u + \frac{v'(x)}{\sqrt{2\alpha}}\right)^2 + 1 + \epsilon_k - \frac{v'(x)^2}{2\alpha} \geq 0$, for $u \in [-1, 1]$.

- Suppose that $\alpha > 2$. Then for any k, $|v_k'(x)| \leq 1 + \frac{\alpha}{2} + \epsilon_k$. Since ϵ_k goes to 0 and $\alpha > 2$, for k sufficiently large, has $\frac{|v_k'(x)|}{\alpha} \leq 1$. Taking $u = -\frac{v_k'(x)}{\alpha}$ gives $1 + \frac{\alpha}{2}u^2 + v_k'(x)u + \epsilon_k \geq 1 + \epsilon_k - \frac{v'(x)^2}{2\alpha} \geq 0$. It must hold μ almost everywhere that $1 + \epsilon_k - \frac{v'(x)^2}{2\alpha} \to 0$ and $|u| = \sqrt{\frac{2}{\alpha}}$.

- Suppose that $0 \leq \alpha \leq 2$. It must hold μ almost every where that $v_k'(x)u \to -1 - \frac{\alpha}{2}u^2$. This implies that $u \neq 0$ and $|v'(x)| \to \frac{1 + \frac{\alpha}{2}u^2}{|u|}$, μ almost every where. It can be verified that for $\alpha \leq 2$, this is a strictly decreasing function of $|u|$ for $|u| \leq 1$. Therefore, it holds that $\liminf |v_k'(x)| \geq 1 + \frac{\alpha}{2}$, μ almost every where. Since we have $|v_k'(x)| \leq 1 + \frac{\alpha}{2} + \epsilon_k$, it holds that $|v_k'(x)| \to 1 + \frac{\alpha}{2}$ and therefore $|u| = 1$, μ almost everywhere. In this case, it is easy to construct alternative sequences $\tilde{v}_k = \frac{2+\tilde{\alpha}}{2+\tilde{\alpha}}v_k$ for $0 \leq \tilde{\alpha} \leq 2$ to show that $l_{\tilde{\alpha}}$ is also a member of $\mathrm{IOCP}(\mu, \mu_T)$.

To conclude, we are in case 1 when the trajectories that generate μ are not optimal with respect to any Lagrangian in $\mathcal{F}$, in particular when $|u|$ is not μ almost everywhere constant. If this is not the case and μ is not degenerate, we have a unique solution or a set of solutions depending on μ and its relation with the constraint on u.

4 Practical inverse control

As discussed in Section 3.2.1, the space of solutions to the inverse problem can be very large. Many of these solutions are of little interest for practitioners because they lack some physical meaning. However, from a formal point of view "valid" solutions exist and ideally they should be the only solutions of a *practical inverse optimal control problem* to be defined.

One may invoke some heuristics to reduce the space of solutions and to enforce prior knowledge in the treatment of the inverse problem. This is commonly achieved by imposing constraints on the candidate Lagrangian solution. Such heuristics include :

- restricting the dependence on certain variables;
- shape conditions (e.g., convexity);
- conic constraints such as positivity;
- parametric constraints (e.g., considering a finite dimensional family of candidate Lagrangians);
- constraints relating the dependence between the candidate Lagrangian and the corresponding value function.

Notice that Definition 2 refers to the large class of continuous Lagrangians with conic constraints. From a theoretical perspective, this allows to characterize inverse optimality in full generality. However this is not amenable to numerical computation yet and so we also describe tractable numerical approximations in the context of inverse optimality.

Finally, according to Definition 2, the input of the inverse problem is an occupation measure. Again, this is a convenient tool for theoretical purposes but in most practical cases such an occupation measure is not available. In fact, roughly speaking, only some realizations of an experiment are available and these realizations form a data set which is an approximation of an hypothetical occupation measure. Therefore in practice the input of the inverse problem is only an approximation of an ideal input, and correctness of this approximation is justified under certain experimental assumptions at the end of this section.

4.1 Normalization

The trivial Lagrangian is solution to the inverse problem independently of the input occupation measures. As we have seen in Section 3.2.1, total variations share the same property. Even though these are solutions to the inverse problem, it is important to avoid them in practice because they do not depend on the input occupation measure and therefore carry no information about it. As illustrated in Example 3.2.3, one way to avoid these spurious solutions is to consider only very restricted families of Lagrangians that cannot contain such solutions. This might be quite restrictive in practice and therefore we provide an alternative. We need the following assumption:

Assumption 2 (Finite time controllability) *There exists $T > 0$ and a compact set $\tilde{X} \subset X$ with smooth boundary and $\mu_0(\tilde{X}) > 0$, such that for any $x_1, x_2 \in \tilde{X}$, there exists $s \in [0, T]$, a bounded function $u : [0, s] \to U$ and an absolutely continuous trajectory $x : [0, s] \to X$ such that $x(0) = x_1$, $x(s) = x_2$ and $\dot{x}(t) = f(x(t), u(t))$ for all $t \in [0, s]$.*

Under assumption 2 we have the following result.

Proposition 1 *If in Definition 2 one includes the normalization*

$$\left| 1 - \int_{\tilde{X} \times U} l + \operatorname{grad} v \cdot f \right| = \epsilon$$

then $0 \notin \mathrm{IOCP}(\mu, \mu_T)$.

Proof : This is due to the following contradiction. Suppose that $0 \in \mathrm{IOCP}(\mu, \mu_T)$. Choose a decreasing sequence $\epsilon_k \to 0$ as $k \to \infty$, and construct a sequence v_k of differentiable functions that satisfy conditions of Definition 2 for the chosen ϵ_k. Then $\langle \mu_0, v_k \rangle = \langle \mu_T, v_k \rangle - \langle \mu, \operatorname{grad} v_k \cdot f \rangle = O(\epsilon_k)$. Because of Assumption 1, $v_k(x_0) \leq T_M \epsilon_k$ for every $x_0 \in X_0$. Therefore, by integration, $\max(\int_{\tilde{X}} v_k \mu_0, \int_{X \setminus \tilde{X}} v_k \mu_0) \leq O(\epsilon_k)$. Furthermore, $\int_{\tilde{X}} v_k \mu_0 + \int_{X \setminus \tilde{X}} v_k \mu_0 = \langle \mu_0, v_k \rangle = O(\epsilon_k)$. Finally,

$$\max \left(\int_{\tilde{X}} v_k \mu_0, - \int_{\tilde{X}} v_k \mu_0 \right) \leq O(\epsilon_k)$$

and $\lim_{k \to \infty} \int_{\tilde{X}} v_k = 0$.

In addition, by Assumption 2, we also have $|v_k(x_1) - v_k(x_2)| \leq T \epsilon_k$ for every $x_1, x_2 \in \tilde{X}$. Since $\mu_0(\tilde{X}) > 0$, this implies that $v_k \to 0$ uniformly on $\tilde{X}$, as $k \to \infty$. In addition, $\int_{\tilde{X} \times U} \operatorname{grad} v_k \cdot f \to 1$ as $k \to \infty$. Next, with the polynomial vector field

$$x \mapsto g(x) = \int_U f(x, u) du,$$

Stokes' Theorem yields

$$\begin{aligned} \int_{\tilde{X}} \operatorname{div}(v_k(x) g(x)) &= \int_{\tilde{X}} \operatorname{div}(g(x)) v_k(x) + \operatorname{grad} v_k(x) \cdot g(x) \\ &= \int_{\tilde{X}} \operatorname{div}(g(x)) v_k(x) + \int_{\tilde{X} \times U} \operatorname{grad} v_k(x) \cdot f(x, u) \\ &= \int_{\partial \tilde{X}} v_k(x) g(x) \cdot \vec{n}(x) \end{aligned}$$

where $\vec{n}(x)$ is the outward pointing normal to the boundary at x. Because of the uniform convergence of $v_k \to 0$ on $\tilde{X}$ as $k \to \infty$, and boundedness of g and $\operatorname{div} g$ on X, the left-hand side converges to 1 while the right-hand side converges to 0, which is a contradiction. □

Remark 5 The conditions on $\tilde{X}$ may be relaxed. Indeed, the only important point is to be able to apply Stokes's Theorem. In particular, the set $\tilde{X}$ could be a box or an open set whose boundary does not have too many non-smooth points, see e.g. [37, Theorem III.14A].

Remark 6 The normalization given in Proposition 1 obviously ensures that Lagrangians in the form of a total variation are excluded.

Assumption 2 may look very strong regarding the result of Proposition 1. However the next example shows that it cannot be excluded.

Example 1 *Consider the direct control problem with $X = U = [0,1]$, $X_T = \{1\}$, $T_M = 1$ and $f(x,u) = u$. These data are obviously not compatible with Assumption 2. Choose $\mu_0(dx) = dx$ and $l_0(x,u) = 1$ so that the couple $\mu(dx,du) := xdx\delta_1(du)$ and $\mu_T(dx) := \delta_1(dx)$, solves the problem*

$$\begin{aligned} p^*(\mu_0) := \quad & \inf_{\mu,\mu_T} \ \langle \mu, 1 \rangle \\ & \text{s.t.} \ \ \mathrm{div} f\mu + \mu_T = \mu_0, \\ & \qquad \langle \mu, 1 \rangle \leq T_M, \\ & \qquad \mu \in \mathcal{M}_+(X \times U), \\ & \qquad \mu_T \in \mathcal{M}_+(X_T). \end{aligned} \tag{7}$$

Indeed, for any differentiable v, $\langle \mathrm{div} f\mu, v \rangle = \int_T^1 -v' f\mu(dx) = \int_T^1 vdx - v(1) = \langle \mu_0, v \rangle - \langle \mu_T, v \rangle$. Furthermore, the function $x \mapsto v^*(x) := 1 - x$ ensures that (μ, μ_T) is an optimal solution. Indeed $\langle \mu, 1 \rangle = \langle \mu_0, v^* \rangle$. Consider a sequence of differentiable functions (v_k), $k \in \mathbb{N}$, such that $v_k(x) = 0$ for $x \geq \frac{1}{k}$ and $v_k(x) = -(kx-1)^2$ otherwise. For $\tilde{X} = [0,1]$, we have $0 \leq \langle \mu, v_k' f \rangle = \int_T^{\frac{1}{k}} kx - k^2x^2 dx = \frac{1}{6k} \to 0$, $v_k' f \geq 0$, $\langle \mu_T, v_k \rangle = 0$, $v_k(1) = 0$, $\int_{\tilde{X} \times U} v_k' f = v_k(1) - v_k(0) = 1$. Therefore even if we enforce the normalization $\int_{\tilde{X} \times U} l + v' f = 1$, we cannot prevent the trivial Lagrangian $l = 0$ from being an optimal solution to the inverse problem.

Remark 7 Regarding Proposition 1, one could argue that simple linear constraints such as $l(0) = 1$ or conic constraints such as $l > 0$ would be sufficient to avoid the trivial Lagrangian. However, this does not allow to avoid total variations which are equivalent to the trivial Lagrangian in terms of solutions to the direct problem.

Denote by NIOCP the subset of IOCP with the normalization constraint of Proposition 1 added to the constraints of Definition 2. One important feature of this normalization is that it can be thought of as a way to intersect the cone of solutions to the inverse problem with an affine subspace. Therefore NIOCP is still closed and convex. Furthermore, if we restrict the set of candidate Lagrangians to be finite-dimensional, one may look for minimum norm-like solutions which will prove to be useful in numerical experiments. Indeed, NIOCP being closed and convex and all norms being equivalent in finite dimensions, optimization problems over NIOCP, if bounded, have an optimal solution. Finally, as $0 \notin$ NIOCP, one may minimize any norm-like function to enforce specific prior structure and avoid the trivial Lagrangian.

4.2 Polynomial approximation

Until now, all the results that we have presented involve continuous and differentiable functions, in full generality. However, for practical computation one has to approximate such functions and of course, *polynomials* are obvious natural candidates. But in our context they are also of particular interest for mainly three reasons:

- for fixed degree, polynomials belong to finite-dimensional spaces and are therefore amenable to computation;

- when varying the degree, the class of polynomials is rich enough to approximate a wide class of functions;
- *Positivity Certificates* from real algebraic geometry allow to express positivity constraints in a computationally tractable way.

From now on, we make the following assumption:

Assumption 3 *f is a polynomial and X, X_T and U are basic semi-algebraic sets.*

As proposed in Definition 2, checking that a polynomial is a solution of the inverse problem involves the construction of a sequence of continuously differentiable functions. These functions can also be approximated by polynomials. In this section we describe some tools required for such an approximation and we also prove the correctness of the approximations.

Let $g_1, \ldots, g_m \in \mathbb{R}[z]$ be polynomials in the variable $z \in \mathbb{R}^n$ and consider the basic semi-algebraic set

$$G = \{z \in \mathbb{R}^n : g_i(z) \geq 0,\, i = 1, \ldots, m\}.$$

Let $g_0 = 1$ and let $\Sigma^2 \subset \mathbb{R}[z]$ denotes the set of sums-of-squares (SOS) polynomials, i.e., $p \in \Sigma^2$ if it can be written as a sum of squares of other polynomials.

Definition 3 *Let $Q_k(G)$ denote the convex cone of polynomials that can be written as*

$$p = \sum_{i=0}^{m} s_i g_i, \quad s_i \in \Sigma^2,\, i = 0, 1, \ldots, m,$$

where the degree of $s_i g_i$, $i = 0, 1, \ldots m$, is at most $2k$. If $p \in Q_k(G)$ we say that p has a Putinar positivity certificate.

It is immediate to check that any element of $Q_k(G)$ is non-negative on G. A remarkable property of such certificates is that a partial converse is true.

Proposition 2 ([29]) *Suppose that the polynomial super-level set $\{x : g_i(x) \geq 0\}$ is compact for some $i = 1, \ldots, m$. If $p > 0$ on G then there exists $k \geq 0$ such that $p \in Q_k(G)$.*

Furthermore and importantly from a computational viewpoint, checking whether $p \in \Sigma^2$ reduces to checking whether a set of LMIs [20] involving the coefficients of p has a solution. Therefore a more precise definition of inverse optimality in the context of polynomial Lagrangians is as follows (compare with Definition 2).

Definition 4 (polyIOCP and polyIOCP$_{\epsilon,k}$) *For $\epsilon > 0$, given measures $\mu \in \mathcal{M}_+(C \times U)$ and $\mu_T \in \mathcal{M}_+(X_T)$ such that $\operatorname{div} f\mu + \mu_T = \mu_0 \in \mathcal{M}_+(X_0)$, denote by $\operatorname{polyIOCP}_{\epsilon,k}(\mu, \mu_T)$ the set of polynomials $l \in \mathbb{R}[x, u]_{2k}$ (i.e., of degree at most $2k$) such that:*

$$\begin{array}{rcl} \langle \mu, l + \operatorname{grad} v \cdot f \rangle & \leq & \epsilon, \\ l + \operatorname{grad} v \cdot f + \epsilon & \in & Q_k(X \times U), \\ \langle \mu_T, v \rangle & \geq & -\epsilon, \\ -v & \in & Q_k(X_T), \end{array}$$

for some polynomial $v \in \mathbb{R}[x]_{2k}$.

Denote also by $\text{polyIOCP}(\mu, \mu_T)$ *the set of polynomial solutions to the inverse optimal control problem: That is,* $l \in \text{polyIOCP}(\mu, \mu_T)$ *if for any* $\epsilon > 0$ *there exists* $k(\epsilon)$ *such that* $l \in \text{polyIOCP}_{\epsilon,k(\epsilon)}$.

In other words, $\text{polyIOCP}_{\epsilon,k}(\mu, \mu_T)$ is the set of polynomial ϵ-solutions with degree bound $2k$, to the inverse optimal control problem. The advantage of the previous definition, is that, provided that one has the possibility to compute the linear functionals $\mu \in \mathbb{R}[x,u]^*$ and $\mu_T \in \mathbb{R}[x]^*$, checking whether $l \in \text{polyIOCP}_{\epsilon,k}(\mu, \mu_T)$ for k and ϵ given reduces to solving a convex LMI problem [20]. Furthermore, under a compactness assumption, in the asymptotic regime this definition is equivalent to Definition 2.

Proposition 3 (Correctness of polynomial approximation) *Suppose that one of the polynomials defining the basic semi-algebraic set* X *(resp.* U*) has a compact super-level set. Then*

$$\text{polyIOCP}(\mu, \mu_T) = \text{IOCP}(\mu, \mu_T) \cap \mathbb{R}[x,u].$$

Proof : The direct inclusion is trivial. For the reverse inclusion, suppose that $l \in \text{IOCP}(\mu, \mu_T) \cap \mathbb{R}[x,u]$. Fix $\epsilon > 0$, and take for v a certificate that $l \in \text{IOCP}_{\frac{\epsilon}{3}}(\mu, \mu_T)$ as given by Definition 2. Since we consider compact sets in finite dimensional spaces, both v and its gradient $\operatorname{grad} v$ can be simultaneously approximated uniformly by a polynomial up to an arbitrary precision. Therefore as f is bounded on $X \times U$, there exists a polynomial v_k of degree k such that $\sup_X |v - v_k| \leq \frac{\epsilon}{3}$ and $\sup_{X\times U} |\operatorname{grad} v \cdot f - \operatorname{grad} v_k \cdot f| \leq \frac{\epsilon}{3}$. Hence $l + \operatorname{grad} v_k \cdot f + \epsilon \geq \frac{\epsilon}{3} > 0$ on $X \times U$ and $\frac{\epsilon}{3} - v_k \geq \frac{\epsilon}{3} > 0$ on X_T. Using Proposition 2, there exists k_1 and k_2 such that $l + \operatorname{grad} v_k \cdot f + \epsilon \in Q_{k_1}(X \times U)$ and $\frac{\epsilon}{3} - v_k \in Q_{k_2}(X_T)$. Then $l \in \text{polyIOCP}_{\epsilon,k}(\mu, \mu_T)$ whenever $k \geq \max(k, k_1, k_2)$ and finally, $l \in \text{polyIOCP}(\mu, \mu_T)$ because ϵ was arbitrary (fixed). □

All properties of Lagrangians in Definition 2 hold for the Lagrangians in Definition 4. For example, as stated in Remark 2, if $l \in \text{polyIOCP}_{\epsilon,k}(\mu, \mu_T)$, then (μ, μ_T) is close to optimal for $\text{OCP}(l, \mu_0, T)$. In particular, if we the moments of μ and μ_T are available then the latter property can be checked numerically by solving a semi-definite program.

4.3 Integral discretization

For practical numerical computation in the context of Definition 4, we still must be able to integrate polynomials with respect to μ and μ_T. This is easy provided that we know the moments of μ and μ_T. However, exact computation of such moments cane be complicated in practice, especially when μ is a superposition of trajectories. Usually, data sets from experiments consist of samples of trajectories which can be seen as realizations of a random sampling process.

In this section we first describe how the framework of occupation measures can formally describe the process of sampling trajectories and we justify the replacement of measures (μ, μ_T) by their empirical counterparts when considering empirical samples as input data for inverse control problems. In the context of polynomial certificates in Definition 4, this

amounts to replacing the moments of the measures (μ, μ_T) by their empirical counterparts. Consider the probability measure μ_0 on X_0 and the measures (μ, μ_T) in equations (1) and (2). One way to interpret these measures is to consider the following random process:

- choose $z \in \text{spt}\, \mu_0$ randomly,
- choose t randomly uniformly on $[0, T_z]$,
- output $\xi = (x_z(t), u_z(t))$.

This defines a generative process for the random variable ξ. If the probability for an initial condition z to belong to a Borel set $A \subset X_0$ is given by

$$\begin{aligned} &\mathbb{P}[z \in A] \\ =&p_0(A) := \frac{\int_A T_z \mu_0(dz)}{\int_{X_0} T_z \mu_0(dz)}, \end{aligned}$$

then the probability for a trajectory ξ to belong to a Borel hyperrectangle $(A, B) \subset X \times U$ is given by

$$\begin{aligned} &\mathbb{P}[\xi \in A \times B] \\ =&p_\mu(A \times B) := \int_{X_0} \left(\int_0^{T_z} I(x_z(t) \in A) I(u_z(t) \in B) dt \right) \mu_0(dz) = \frac{\mu(A \times B)}{\mu(X \times U)}. \end{aligned}$$

A statistical model for points (x_i, u_i), $i = 1, \ldots, n$, that are samples of trajectories, is to assume that we repeat the previous process n times, independently. In this case, we say that the database $\mathcal{D} = \{(x_i, u_i)\}_{i=1}^n$ is made of independent realizations of a random variable with underlying distribution p_μ. The process which generates the database being random, we write

$$\{\xi_i\}_{i=1}^n \overset{i.i.d.}{\sim} p_\mu.$$

to stress that all ξ_i are independent and identically distributed (*i.i.d.*) according to p_μ (they are independent copies of the same random variable). Similarly, μ_T can be seen as the probability distribution describing the following random process:

- choose $z \in \text{spt}\, \mu_0$ randomly according to p_0,
- output $x_z(T_z)$.

We now define what is an approximate solution to the inverse problem when the only information available about μ is a realization, $\mathcal{D} = \{(x_i, u_i)\}_{i=1}^n$, of a random process, $\{\xi_i\}_{i=1}^n \overset{i.i.d.}{\sim} p_\mu$.

Definition 5 (Sampled-IOCP$_{\epsilon,k}$) *For $\epsilon > 0$ and $\mathcal{D} = \{(x_i, u_i)\}_{i=1}^n$, let* Sampled $-$ IOCP$_{\epsilon,k}$ *be the set of polynomials $l \in \mathbb{R}[x,u]_{2k}$ such that :*

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} l(x_i,u_i) + \operatorname{grad} v(x_i)\cdot f(x_i,u_i) \leq \epsilon,\\
&l + \operatorname{grad} v \cdot f \in Q_k(X\times U),\\
&-v \in Q_k(X_T),\\
&v+\epsilon \in Q_k(X_T),\\
&\int_{\tilde{X}\times U} l + \operatorname{grad} v\cdot f = 1.
\end{aligned}$$

for some polynomial $v \in \mathbb{R}[x]_{2k}$. In other words, Sampled $-$ IOCP$_{\epsilon,k}$ *is the set of polynomial ϵ-optimal solutions (with degree bound k) of the sampled inverse optimal control problem.,*

One has replaced μ by its empirical counterpart, added the normalization of Proposition 1 and simplified other conditions; see also Remark 8.

Importantly, membership in Sampled $-$ IOCP$_{\epsilon,k}(\mathcal{D})$ *can be tested by semi-definite programming.*

Using arguments from empirical processes and learning theory, one can quantify the price to pay for this discretization. Of course since we assume that the process that generates the data include some randomness, such a quantification holds probabilitically.

Proposition 4 *Suppose that $\mathcal{D} = \{(x_i,u_i)\}_{i=1}^n$ is a realization of the random process $\{\xi_i\}_{i=1}^n \overset{iid}{\sim} p_\mu$. Then there exist constants $K_1(X,U,k)$, $K_2(X,U,k)$ that only depend on (X,U), and $k \in \mathbb{N}$, such that for any $l \in$* Sampled $-$ IOCP$_{\epsilon,k}(\mathcal{D})$ *and any $0 < \delta < 1$:*

$$l \in \text{polyIOCP}_{\epsilon'(n),k}(\mu,\mu_T) \text{ with probability } 1-\delta,$$

where

$$\epsilon'(n) = \epsilon + \frac{1}{\sqrt{n}}\left(K_1(X,U,k) + K_2(X,U,k)\sqrt{\ln\frac{2}{\delta}}\right),$$

and where the randomness comes from the realization of $\mathcal{D}$.

Proof : Apply Lemma 2 of Appendix A to the polynomial $l + \operatorname{grad} v \cdot f$ to get a bound on $\langle p_\mu, l + \operatorname{grad} v \cdot f\rangle$. □

Remark 8 The conditions detailed in Definition 5 ensure that $v \leq 0$ on X_T and $\langle \mu_T, v\rangle \geq -\epsilon$, for any terminal measure. This is done in order to avoid to deal with the terminal measure μ_T but other alternatives are possible. For example, when X_T is a single point or a simple algebraic set one may enforce (as we do in Section 5) $v = 0$ on X_T instead. Another possibility is to replace μ_T by its empirical counterpart. However in this case we need to provide a lower bound or add constraints on v to obtain finite sample bounds as described in Proposition 4.

Proposition 4 mixes arguments from measure theory and conic optimization with arguments from empirical process and statistical learning theory. The implication of this result is that for a fixed degree, provided that the sample size is big enough, with high probability we do not loose much by approximating μ by an empirical sample.

• We would like to emphasize here that it is necessary to restrict the complexity of the class of functions in which the candidate Lagrangian is searched. Indeed otherwise for instance, for any fixed sample, the polynomial $(x,u) \mapsto \prod_{i=1}^n ||x-x_i||^2||u-u_i||^2$ belongs to $\text{Sampled} - \text{IOCP}_{0,4n}(\mathcal{D})$, but this clearly does not give much insight on the original control problem! The degree of the polynomial candidates is one among many possible measures of complexity. Furthermore, although the constants are likely to be sub-optimal, they give a sense of how fast the degree of the polynomial approximation may grow with respect to the sample size in order to maintain accurate approximations of μ.

5 Numerical illustrations

Building on results of Section 4, we next provide illustrative numerical simulations. In order to fit in the framework of the previous section, we consider examples where f is a polynomial and X, U and X_T are basic semi-algebraic sets. In addition, the input data of the inverse problem is given by a finite database: $\mathcal{D} = \{(x_i, u_i)\}_{i=1,\dots,n}$. In the sequel, the candidate Lagrangians satisfy Definition 5.

To compute such Lagrangians, the main idea is to solve an optimization problem with fixed $k \in \mathbb{N}$, where:

- $l \in \mathbb{R}[x,u]_{2k}, v \in \mathbb{R}[x]_{2k}$ and $\epsilon > 0$ are the decision variables,
- ϵ is the criterion to minimize,
- $\text{Sampled} - \text{IOCP}_{\epsilon,k}(\mathcal{D})$ is the projection on (l,ϵ) of the set of feasible solutions (l,ϵ,v).

In addition, we also include a sparsity inducing term in the criterion that will prove to be useful in numerical experiments.

5.1 Numerical experiments

5.1.1 Problem formulation

We consider the following optimization problem:

$$\begin{aligned} &\inf_{l,v,\epsilon} \ \epsilon + \lambda||l||_1 \\ &\text{s.t. } \tfrac{1}{n}\textstyle\sum_{i=1}^n l(x_i,u_i) + \text{grad}\, v(x_i)\cdot f(x_i,u_i) \le \epsilon, \\ &\quad l + \text{grad}\, v \cdot f \in Q_k(X\times U), \\ &\quad v = 0 \ \text{ on } \ X_T, \\ &\quad \textstyle\int_{\tilde{X}\times U} l + \text{grad}\, v \cdot f = 1. \end{aligned} \tag{iocp}$$

where $l \in \mathbb{R}[x,u]_{2k}$, $v \in \mathbb{R}[x]_{2k}$, ϵ is a real, $\lambda > 0$ (fixed) is a given regularization parameter, and $||.||_1$ denotes the ℓ_1 norm of a polynomial, *i.e.* the sum of absolute values of its coefficients when expanded in the monomial basis. The first constraints come from Definition 5 and the last affine constraint is meant to avoid the trivial solution; see Proposition 1. The ℓ_1 norm is not differentiable around sparse vectors (with entries equal to zero) and has the sparsity promoting role to bias solutions of the problem towards polynomial Lagrangian solutions with few nonzero coefficients.

This regularization affects problem well-posedness and will prove to be essential in numerical experiments.

5.1.2 Numerical implementation

Linear constraints are easily expressed in term of polynomial coefficients. A classical lifting allows to express the ℓ_1 norm as a linear program: for $x \in \mathbb{R}^n$, $||x||_1 = \min_s \sum_{i=1}^n s_i$ subject to $s_i \geq x_i$ and $s_i \geq -x_i$, for all i. The Putinar positivity certificates can be expressed as LMIs [20] whose size depends on the degree bound k. We use the SOS module of the `YALMIP` toolbox [22] to manipulate and express polynomial constraints at a high level in `MATLAB`. The size of the corresponding LMI grows as $\binom{n+k}{k}$ where n is the number of variables and k the degree bound in Putinar certificates. Thus it is reasonable to consider relatively small problems. As shown in the numerical results section, we could handle problems with 5 variables and degree 10 with a reasonable amount of time and memory. To handle larger size problems, specific heuristics and techniques beyond the scope of this paper must be implemented.

5.1.3 General setting

We consider several direct problems of the same form as (ocp_0). That is, we give ourselves compact basic semi-algebraic sets X, U, X_T, the dynamics f, and a Lagrangian l_0. We take known examples for which the (direct) optimal control law can be computed and try to vary their degree of difficulty. Given these optimal state-control trajectories, we generate randomly n data points $\mathcal{D} = \{(x_i, u_i)\}_{i=1\ldots n}$ according to the random process described in Section 5. For a given value of λ and k, we compute a solution l of problem (iocp). Then we measure how l is close to l_0 by computing the following quantity (in the monomial basis):

$$\min_{\alpha} \frac{||l_0 - \alpha l||_2}{||l_0||_2} = \left(1 - \frac{\langle l_0, l\rangle^2}{||l_0||_2^2 ||l||_2^2}\right)^{\frac{1}{2}}. \tag{8}$$

We also report the value of ϵ in program (iocp). A larger value of ϵ means less reliable numerical certificates; see Remark 2.

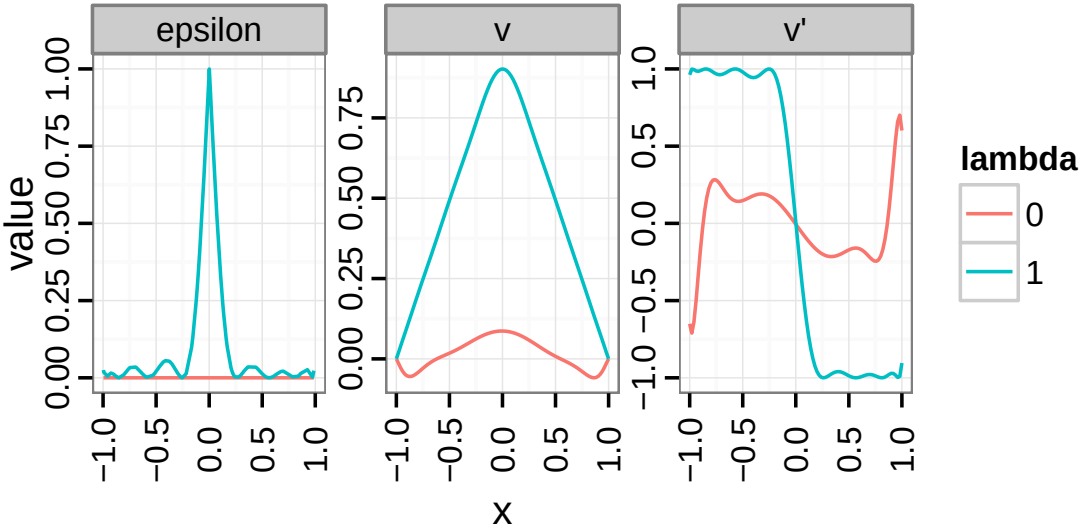

Figure 1: Solution for the one dimensional minimum exit time problem, effect of the regularization parameter λ. See Section 5.2 for problem details and comments. The first column is the distribution of the error ϵ. It represents the value of $l(x,u) + \operatorname{grad} v \cdot f(x,u)$ as a function of x where $u = \operatorname{sign} x$ is the optimal control. The second column is a representation of the value function v and the third column is a representation of its derivative for solutions of problem (iocp) with and without regularization. We take 100 points on the segment. Lagrangian l and value function v are both polynomials of degree 16.

5.2 Illustration on a one-dimensional example

First consider the eikonal problem of minimum exit time from the unit ball in the one-dimensional case. The data of the problem are

$$X = U = B_1,\ X_T = \partial X,\ l_0 = 1,\ f = u.$$

The optimal law for this problem is $u = \operatorname{sign} x$ and the value function is $v_0(x) = 1 - |x|$. We sample 100 points uniformly in X and solve problem (iocp). We compare the choices $\lambda = 0$ (no regularization) and $\lambda = 1$. Results are presented in Figure 1 which displays the distribution of the error ϵ, the estimated value function v as well as its first derivative. Despite the simplicity of the problem, it is quite representative of the difficulties that arise in the context of inverse optimality. The first difficulty is the size of the set of solutions to the inverse problem:

- given any symmetric differentiable concave function v vanishing on $\{-1, 1\}$, the pair $(l = |v'|, v)$ solves problem (iocp) with $\epsilon = 0$;
- any positive polynomial on $X \times U$ vanishing if $|u| = 1$ solves problem (iocp) with $\epsilon = 0$;
- any Lagrangian of the form $v'(x)u$ solves problem (iocp) with $\epsilon = 0$;

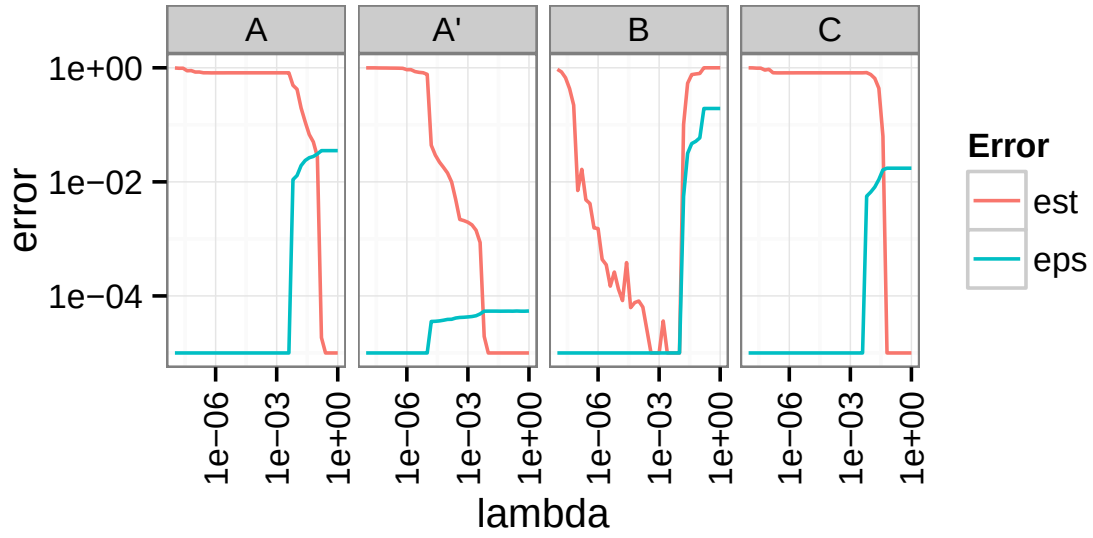

Figure 2: Error versus regularization parameter λ. Problems details are given in the main text. Estimation error (est) is given in (8). Epsilon error (eps) is the value of ϵ in program (iocp). Trajectory sample size: 20 for A, A' and B, 50 for C. Degrees of l and v are 4 and 10 respectively.

- any convex combination of solutions of the types mentioned above also solves problem (iocp).

Even though formally accurate, these solutions form a relatively large set and do not carry any physical meaning. In the absence of any additional form of prior knowledge, it is impossible to discriminate between these solutions and the one that we wish to recover, namely l_0. This is illustrated in Figure 1 where the red line ($\lambda = 0$) displays an example of value function v obtained with a very low value of ϵ. This is a very good certificate that our database $\mathcal{D}$ is close to optimal for the corresponding Lagrangian l. However, the estimated Lagrangian is far from the original one, namely $l_0 = 1$. Moreover, the shape of the value function is quite uncommon. This motivates the use of prior knowledge to bias the solutions of problem (iocp) toward a certain set of solutions. We use the ℓ_1 norm which tends to promote Lagrangians with few non-zero coefficients.

When $\lambda = 1$, the sparsity inducing effect of ℓ_1-norm regularization allows to recover the true Lagrangian ($l_0 = 1$) which is indeed sparse. The solution of problem (iocp) involves a polynomial function v which should in principle be close to the true value function $v_0(x) = 1 - |x|$. The v function displayed in Figure 1 is close to v_0. However, v_0 is not smooth around the origin and therefore its derivative is harder to approximate by polynomials around this point. Hence the value of the error is higher around the origin.

5.3 Illustration on more complex problems

These simulations are taken from [27]. We consider the following free terminal time direct problems:

Minimum exit time in dimension 2:

$$X = U = B_2,\ X_T = \partial X,\ l_0 = 1,\ f = u. \tag{9}$$

The optimal law is $u = \frac{x}{||x||_2}$ and the value function is $v_0(x) = 1 - ||x||_2$.

Minimum exit norm in dimension 2:

$$X = U = B_2,\ X_T = \partial X, l_0 = ||x||_2^2 + ||u||_2^2,\ f = u. \tag{10}$$

The optimal law is $u = x$ and the value function is $v_0(x) = 1 - ||x||_2^2$.

Minimum time Brockett integrator:

$$X = 3B_3,\ U = B_2,\ l_0 = 1,\ f = (u_1, u_2, u_1x_2 - u_2x_1). \tag{11}$$

Recall that the Brockett integrator of nonlinear systems control is also known (up to a change of coordinates) as the unicycle or Dubins system, one of the simplest instance of a non-holonomic system in robotics, see e.g. [12] for the connection. The optimal law and value function are described in [28]. Complementary details are found in Appendix B.

Data generation: We consider the following settings:

A problem (9) with samples from B_2;

A' problem (9) with samples from $B_2 \setminus \frac{1}{2}B_2$;

B problem (10) with samples from B_2;

C problem (11) with samples from $B_3 \setminus \left(\frac{1}{2}B_2 \times \mathbb{R}\right)$.

In all cases we fix the degree of l to 4 and that of v to 10.

Results: The results for the four problems are presented in Figure 2. For all problems, l is of degree 4. Therefore, l is to be found in a space of dimension 70 for problems A, A' and 126 for problem D. When the estimation error is close to 1, we estimate a Lagrangian l that is orthogonal to l_0 (in the monomial basis), and when it is close to 0, they are colinear. We also display the ϵ value of (iocp). We consider that the estimation is reasonable, when both estimation error and ϵ values are low.

• For all problems we are able to recover the true Lagrangian with good accuracy for some value of the regularization parameter λ. In the absence of regularization, we do not recover the true Lagrangian at all. This highlights the important role of ℓ_1 regularization which allows to bias the estimation toward *sparse polynomials*. The choice of λ in practical settings is subject to heuristics: numerical simulations or cross-validation which consists in keeping a portion of the input data as a validation set.

• For all four problems, when the estimation error is minimal, the value of ϵ is reasonably low, depending on how the value function can be approximated by a polynomial. For example, A' shows lower ϵ value because we avoid sampling database points close to the non-differentiable point of the true value function. In example D, the value function is known to be harder to approximate by polynomials and the value of ϵ is a bit larger. The estimation accuracy is still very reasonable.

6 Conclusion

The main contribution of this paper is to propose a general framework to analyze the inverse problem of optimal control. The analysis is based on the weak formulation of direct optimal control problems using occupation measures, relaxed Hamilton-Jacobi-Bellman optimality conditions, and duality of infinite-dimensional linear programs. The proposed formulation is powerful enough to ensure that there is no gap between solutions of the direct and inverse problem (Theorem 1). To the best of our knowledge this is the first result of this kind. In addition, in principle the proposed methodology is applicable to practical problems where we only have access to sample trajectories. We have also proposed numerical and statistical approximation procedures from which solid theoretical guaranties can be obtained. Finally we have illustrated our results on relatively simple (but not trivial) numerical examples of modest size.

One of the most striking aspects of the inverse problem is its set of valid solutions. Indeed, even for the simplest problems it is difficult to discriminate between physically meaningful Lagrangians and spurious mathematical solutions. For this reason, formulating the inverse problem as a well-posed problem (in particular with a unique solution) requires the introduction of strong prior knowledge – sometimes arguably too restrictive – about the nature of the Lagrangian to be recovered; see for example [2]. However the proposed formulation based on relaxed HJB-optimality conditions, allows to get intuitions about characteristics that affect well-posedness of the problem.

This work is to be seen as a first step toward a theoretical and practical framework for the resolution of inverse problems in a variety of contexts. Further aspects of the problem have to be investigated within this realm. First, we only deal with deterministic trajectories. For practical purposes it is essential to consider the effect of experimental noise, both from theoretical and practical perspectives, and to determine to which extent and how the problem can be solved in this more difficult context. Second, we have proposed a numerical scheme to approximate solutions and show that it is effective on academic examples of modest size. Experimental validation of such approximations should be carried out on real world examples of larger size. This involves a lot of data processing and fine tuning for each specific example. In this perspective, humanoid robotics provides an active and attractive field of application [4, 24].

Acknowledgments

This work was partly funded by an award of the *Simone and Cino del Duca foundation* of Institut de France, a grant of the Gaspard Monge program (PGMO) of the *Fondation mathématique Jacques Hadamard*. Most of this work was carried out during Edouard Pauwels' postdoctoral stay at LAAS-CNRS. The authors would like to thank Frédéric Jean, Jean-Paul Laumond, Nicolas Mansard and Ulysse Serres for fruitful discussions.

A Proof of Proposition 4

We develop uniform finite sample bounds that hold with high probability for arbitrary probability distribution in the context of polynomial functions. These are in particular useful to derive bounds for the random process described by occupation measures as exposed in Section 4.3. The techniques used have become fairly standard in empirical process theory and statistical learning theory, see for example [9] for a nice introduction.

In what follows, we consider a compact set $Z \subset \mathbb{R}^p$ with non empty interior. For a polynomial $z \in \mathbb{R}^p \mapsto q \in \mathbb{R}_d[z]$ of degree d, $c(q)$ denotes its coefficients in the monomial basis (of size $\binom{p+d}{d}$). Similarly for a point $z \in Z$, $v(z)$ denotes the $\binom{p+d}{d}$ dimensional vector representing the evaluation of the corresponding monomials at z such that $q(z) = c(q) \cdot v(z)$ with the dot denoting the inner product. We consider the following set of polynomials $K_d(Z) = \{q \in \mathbb{R}_d[z],\, q \geq 0 \text{ on } Z,\, \int_{\tilde{Z}} q = 1\}$ where $\tilde{Z}$ is a closed subset of Z with nonempty interior. We fix an arbitrary probability distribution P on Z. We denote by ℓ the linear functional on the space $\mathcal{C}(Z)$ such that

$$\langle \ell, f \rangle = \int_Z f(z) P(dx).$$

Similarly for a sample of size n, $S_n = \{z_1, \ldots, z_n\} \in Z^n$, drawn *iid* from P, we denote by ℓ_n the linear functional on the space $\mathcal{C}(Z)$ such that

$$\langle \ell_n, f \rangle = \frac{1}{n} \sum_{i=1}^{n} f(z_i)$$

for any function f continuous on Z.

Lemma 2 *For any $0 \leq \delta \leq 1$ and any $q \in K_d(Z)$, it holds with probability $1 - \delta$ that*

$$\langle \ell, q \rangle \leq \langle \ell_n, q \rangle + 2\frac{M_c^d(Z) M_v^d(Z)}{\sqrt{n}} + M_\infty^d(Z) \sqrt{\frac{1}{2n} \ln \frac{2}{\delta}},$$

where

$$\begin{aligned} M_\infty^d(Z) &= \sup_{p \in K_d(Z),\, z \in Z} p(z), \\ M_c^d(Z) &= \sup_{p \in K_d(Z)} ||c(p)||, \\ M_v^d(Z) &= \sup_{z \in Z} ||v(z)||_2 \end{aligned}$$

are finite quantities that only depend on X and d.

Proof : The proof combines standard arguments from statistical learning which we describe here for completeness. In the sequel, given a probability distribution P on Z, we use the notation

$$\mathbb{P}[A] = \int_A P(dz_1) \dots P(dz_n), \quad \mathbb{E}_z[F] = \int_A F(z_1, \dots, z_n) P(dz_1) \dots P(dz_n)$$

and we rely on the following concentration result:

Lemma 3 (McDiarmid's inequality [23]) *Assume for all i = 1, ..., n,*

$$\sup_{z_1,\dots,z_n,z_i' \in Z} |F(z_1, \dots, z_i, \dots, z_n) - F(z_1, \dots, z_i', \dots, z_n)| \leq \alpha,$$

then, for all $\epsilon \geq 0$, when $\{z_i\}_{i=1}^n$ is drawn iid from a probability distribution P on Z, we have

$$\mathbb{P}\left[|F - \mathbb{E}_z[F]| \geq \epsilon\right] \leq 2 \exp\left(-\frac{2\epsilon^2}{n\alpha^2}\right),$$

where the expectation is taken over the random sample. An equivalent formulation is that for $0 \leq \delta \leq 1$, with probability $1 - \delta$, it holds that

$$F \leq \mathbb{E}_z[F] + \alpha \sqrt{\frac{n}{2} \ln \frac{2}{\delta}}.$$

We consider the following quantity

$$F(z_1, \dots, z_n) := \sup_{p \in K_d(Z)} \langle \ell - \ell_n, p \rangle .$$

Observe that $K_d(Z)$ is a subset of a finite dimensional space and that for $p \in K_d(Z)$, we have $||p||_1^{\tilde{Z}} = \int_{\tilde{Z}} |p| = 1$. Since all norms are equivalent, $K_d(Z)$ is bounded in any given norm on polynomials, in particular, the supremum norm. Therefore, the quantity

$$M_\infty^d(Z) := \sup_{p \in K_d(Z)} ||p||_\infty^{\tilde{Z}} := \sup_{p \in K_d(Z), z \in Z} p(z)$$

is finite. We have that for all i, and any $z_1, \dots, z_n, z_i'$ and any $p \in K_d(Z)$

$$|F(z_1, \dots, z_i, \dots, z_n) - F(z_1, \dots, z_i', \dots, z_n)| \leq \sup_{p \in K_d(Z)} \frac{1}{n} |p(z_i) - p(z_i')| \leq \frac{M_\infty^d(Z)}{n}.$$

Therefore McDiarmid's inequality of Lemma 3 applies to function F with $\alpha = \frac{M_\infty^d(X)}{n}$, and, for any $q \in K_d(Z)$, with probability $1 - \delta$, it holds that

$$\langle \ell - \ell_n, q \rangle \leq \sup_{p \in K_d(Z)} \langle \ell - \ell_n, p \rangle \leq \mathbb{E}_z \left[\sup_{p \in K_d(Z)} \langle \ell - \ell_n, p \rangle \right] + M_\infty^d(Z) \sqrt{\frac{1}{2n} \ln \frac{2}{\delta}}. \quad (12)$$

The left hand side depends on the random draw of the sample $\{z_i\}_{i=1}^n$, but the right hand side is deterministic. We use a standard symmetrization argument to bound the

expectation in the right hand side. Using the definition of ℓ and ℓ_n, the convexity of the supremum and Jensen's inequality, we have that

$$\begin{aligned}\mathbb{E}_z\left[\sup_{p\in K_d(Z)}\langle \ell-\ell_n,p\rangle\right] &= \mathbb{E}_z\left[\sup_{p\in K_d(Z)}\mathbb{E}_{z'}\left[\langle \ell'_n,p\rangle\right]-\langle \ell_n,p\rangle\right] \\ &\leq \mathbb{E}_{z,z'}\left[\sup_{p\in K_d(Z)}\langle \ell'_n-\ell_n,p\rangle\right] \\ &= \mathbb{E}_{z,z'}\left[\sup_{p\in K_d(Z)}\frac{1}{n}\sum_{i=1}^n p(z'_i)-p(z_i)\right]\end{aligned} \tag{13}$$

where the notation z' refers to any other sample $S'=\{z'_i\}_{i=1}^n$ drawn from P and $\ell_{n'}$ is the corresponding empirical measure. The *iid* assumption allows to flip z_i and z'_i in the expectation. Let ξ_i be Rademacher variables, i.e. random variables which take values in $\{-1,1\}$, each with probability one half. We have

$$\begin{aligned}&\mathbb{E}_{z,z'}\left[\sup_{p\in K_d(Z)}\frac{1}{n}\sum_{i=1}^n p(z'_i)-p(z_i)\right] \\ =&\mathbb{E}_{z,z',\xi}\left[\sup_{p\in K_d(Z)}\frac{1}{n}\sum_{i=1}^n \xi_i(p(z'_i)-p(z_i))\right] \\ \leq&\mathbb{E}_{z,z',\xi}\left[\sup_{p\in K_d(Z)}\frac{1}{n}\sum_{i=1}^n \xi_i p(z'_i)+\sup_{p\in K_d(Z)}\frac{1}{n}\sum_{i=1}^n -\xi_i p(z_i)\right] \\ =&2\mathbb{E}_{z,\xi}\left[\sup_{p\in K_d(Z)}\frac{1}{n}\sum_{i=1}^n \xi_i p(z_i)\right] \\ =&2\mathbb{E}_{z,\xi}\left[\sup_{p\in K_d(Z)}\frac{1}{n}c(p)\cdot\sum_{i=1}^n \xi_i v(z_i)\right].\end{aligned} \tag{14}$$

The quantity on the right hand side is known as the Rademacher complexity of the function class $K_d(Z)$. Intuitively, it measures to which extent elements of a function class correlate with random noise in a worst case scenario. The function $p\to||c(p)||_2$ is a norm on polynomials and since $K_d(Z)$ is bounded, the quantity

$$M_c^d(Z):=\sup_{p\in K_d(Z)}||c(p)||_2$$

is finite. Moreover, since Z is compact, the quantity

$$M_v^d(Z):=\sup_{z\in Z}||v(z)||_2$$

is also finite and attained. We have that

$$\begin{aligned}\sup_{p\in K_d(Z)}\frac{1}{n}c(p)\cdot\sum_{i=1}^n \xi_i v(z_i) \leq&\frac{M_c^d(Z)}{n}\left|\left|\sum_{i=1}^n \xi_i v(z_i)\right|\right|_2 \\ =&\frac{M_c^d(Z)}{n}\sqrt{\sum_{i=1}^n\sum_{j=1}^n \xi_i\xi_j v(z_i)\cdot v(z_j)}.\end{aligned}$$

¡++¿ Moreover, $\mathbb{E}_\xi\left[\sum_{j=1}^n \xi_i\xi_j v(z_i)\cdot v(z_j)\right] = \sum_{i=1}^n v(z_i)^2$ ($\mathbb{E}_\xi[\xi_i\xi_j] = I(i=j)$). Therefore, using Jensen's inequality (with concavity of the square root), we obtain

$$\mathbb{E}_\xi\left[\sup_{p\in\mathcal{K}} \frac{1}{n} c(p)\cdot \sum_{i=1}^n \xi_i v(x_i)\right] \leq \frac{M_c^d(X)}{n}\sqrt{\sum_{i=1}^n ||v(x_i)||_2^2} \leq \frac{M_c^d(X)M_v^d(X)}{\sqrt{n}}.$$

Putting things together, using inequalities (12), (13), (14), we have that with probability $1-\delta$, it holds

$$\langle c, q\rangle \leq \langle c_n, q\rangle + 2\frac{M_c^d(X)M_v^d(X)}{\sqrt{n}} + M_\infty^d(X)\sqrt{\frac{1}{2n}\ln\frac{2}{\delta}}.$$

□

B Minimal time Brockett integrator

This section covers materials that were presented in [28]. Equation (21) and Lemma 1 in [28] are directly related to the computation of optimal control for problem C in Section 5, but they contain a misprint. We expose here details that have been acknowledged to be correct by the authors of [28]. The free terminal time problem can be formulated as follows:

$$\begin{aligned} T(x_1,x_2,x_3) = \inf_{S,u} \int_0^S dt \qquad & (15)\\ \dot{x}(t) &= \begin{pmatrix} u_1(t) \\ u_2(t) \\ u_1(t)x_2(t) - u_2(t)x_1(t) \end{pmatrix},\\ x(0) &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},\ x(S) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},\\ \|u\|_2 &\leq 1. \end{aligned}$$

If T is differentiable at (x_1,x_2,x_3), then, by Bellman's principle of optimality, one must have

$$\begin{aligned} &\inf_{||u||\leq 1} \operatorname{grad} T(x_1,x_2,x_3)\cdot \begin{pmatrix} u_1 \\ u_2 \\ u_1(t)x_2(t) - u_2(t)x_1(t) \end{pmatrix} + 1 \qquad (16)\\ &= \inf_{\|u\|\leq 1} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\cdot \begin{pmatrix} \frac{\partial T}{\partial x_1} + \frac{\partial T}{\partial x_3}x_2 \\ \frac{\partial T}{\partial x_2} - \frac{\partial T}{\partial x_3}x_1 \end{pmatrix} + 1\\ &= 1 - \left\| \begin{pmatrix} \frac{\partial T}{\partial x_1} + \frac{\partial T}{\partial x_3}x_2 \\ \frac{\partial T}{\partial x_2} - \frac{\partial T}{\partial x_3}x_1 \end{pmatrix}\right\|_2\\ &= 0, \end{aligned}$$

with the optimal feedback given by

$$u = -\begin{pmatrix} \frac{\partial T}{\partial x_1} + \frac{\partial T}{\partial x_3} x_2 \\ \frac{\partial T}{\partial x_2} - \frac{\partial T}{\partial x_3} x_1 \end{pmatrix}. \tag{17}$$

In particular, it must hold that

$$\left\| \begin{pmatrix} \frac{\partial T}{\partial x_1} + \frac{\partial T}{\partial x_3} x_2 \\ \frac{\partial T}{\partial x_2} - \frac{\partial T}{\partial x_3} x_1 \end{pmatrix} \right\|_2 = 1. \tag{18}$$

It is known ([8, Theorems 1.36 and 1.56]) that, provided that $(x_1, x_2) \neq (0, 0)$, T is differentiable and we have

$$T(x_1, x_2, x_3) = \frac{\theta}{\sqrt{\theta + \sin^2\theta - \sin\theta\cos\theta}} \sqrt{x_1^2 + x_2^2 + 2|x_3|} \tag{19}$$

where θ is the unique solution in $[0, \pi)$ to

$$g(\theta) = \frac{2|x_3|}{x_1^2 + x_2^2} \tag{20}$$

with

$$g(\theta) = \frac{\theta - \sin\theta\cos\theta}{\sin^2\theta}. \tag{21}$$

Let us compute the gradient of T to get an expression of the optimal control as in (17). From (20), we get

$$\frac{\partial\theta}{\partial x_i} g'(\theta)(x_1^2 + x_2^2) + 2g(\theta)x_i = 0,$$

for $i = 1, 2$. This leads to

$$\frac{\partial\theta}{\partial x_i} = -\frac{2g(\theta)x_i}{g'(\theta)(x_1^2 + x_2^2)},$$

for $i = 1, 2$. Similarly, we have (assume that $x_3 \neq 0$)

$$\frac{\partial\theta}{\partial x_3} = \frac{2\,\mathrm{sign}\, x_3}{g'(\theta)(x_1^2 + x_2^2)}.$$

Moreover it holds that

$$\begin{aligned} T(x_1, x_2, x_3) &= \frac{\theta}{\sin\theta} \sqrt{x_1^2 + x_2^2} \\ g'(\theta) &= 2\frac{\sin\theta - \theta\cos\theta}{\sin^3\theta} \\ \frac{\theta}{\sin\theta} &= g(\theta)\sin\theta + \cos\theta \\ \frac{\partial}{\partial\theta}\left(\frac{\theta}{\sin\theta}\right) &= \frac{\sin\theta - \theta\cos\theta}{\sin^2\theta} = \frac{1}{2}\sin\theta g'(\theta). \end{aligned} \tag{22}$$

We have, for $i = 1, 2$,

$$\begin{aligned}\frac{\partial T}{\partial x_i} &= \frac{\partial \theta}{\partial x_i}\frac{\partial}{\partial \theta}\left(\frac{\theta}{\sin\theta}\right)\sqrt{x_1^2+x_2^2} + \frac{\theta}{\sin\theta}\frac{x_i}{\sqrt{x_1^2+x_2^2}} \\ &= \frac{-2g(\theta)x_i}{g'(\theta)(x_1^2+x_2^2)}\frac{\sin\theta g'(\theta)}{2}\sqrt{x_1^2+x_2^2} + \frac{g(\theta)\sin\theta x_i + \cos\theta x_i}{\sqrt{x_1^2+x_2^2}} \\ &= \frac{\cos\theta x_i}{\sqrt{x_1^2+x_2^2}}.\end{aligned} \tag{23}$$

Moreover

$$\begin{aligned}\frac{\partial T}{\partial x_3} &= \frac{\partial \theta}{\partial x_3}\frac{\partial}{\partial \theta}\left(\frac{\theta}{\sin\theta}\right)\sqrt{x_1^2+x_2^2} \\ &= \frac{2\ \mathrm{sign}\ x_3}{g'(\theta)(x_1^2+x_2^2)}\frac{\sin\theta g'(\theta)}{2}\sqrt{x_1^2+x_2^2} \\ &= \frac{\mathrm{sign}\ x_3\ \sin\theta}{\sqrt{x_1^2+x_2^2}}.\end{aligned} \tag{24}$$

It can be verified that equation (18) holds for these values of the partial derivatives. The optimal control can be computed by combining (17) , (23) and (24).

References

[1] P. Abbeel and A. Y. Ng. Apprenticeship learning via inverse reinforcement learning. Proceedings of the International Conference on Machine Learning, ACM, 2004.

[2] A. Ajami, J. P. Gauthier, T. Maillot and U. Serres. How humans fly. ESAIM: Control, Optimisation and Calculus of Variations, 19(4):1030–1054, 2013.

[3] B. D. O. Anderson and J.B. Moore. Linear optimal control. Prentice-Hall, Englewood Cliffs, NJ, 1971.

[4] G. Arechavaleta, J. P. Laumond, H. Hicheur and A. Berthoz. An optimality principle governing human walking. IEEE Transactions on Robotics, 24(1):5–14, 2008.

[5] M. Athans and P. L. Falb. Optimal control. An introduction to the theory and its applications. McGraw-Hill, New York,1966.

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

[7] A. Barvinok. A course in convexity. AMS, Providence, NJ, 2002.

[8] R. Beals, B. Gaveau, and P. C. Greiner. Hamilton-Jacobi theory and the heat kernel on Heisenberg groups. Journal de Mathématiques Pures et Appliquées, 79(7):633–689, 2000.

[9] O. Bousquet, S. Boucheron, and G. Lugosi. Introduction to statistical learning theory. In Advanced Lectures on Machine Learning, 169–207, Springer, Berlin, 2004.

[10] J. Casti. On the general inverse problem of optimal control theory. Journal of Optimization Theory and Applications, 32(4):491–497, 1980.

[11] F. C. Chittaro, F. Jean, and 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.

[12] D. DeVon and T. Bretl. Kinematic and dynamic control of a wheeled mobile robot IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007.

[13] R.A. Freeman and P.V. Kokotović. Inverse optimality in robust stabilization. SIAM Journal on Control and Optimization, 34(4):1365–1391, 1996.

[14] T. Fujii and M. Narazaki. A complete optimality condition in the inverse problem of optimal control. SIAM Journal on Control and Optimization, 22(2):327–341, 1984.

[15] V. Gaitsgory and M. Quincampoix. Linear Programming Approach to Deterministic Infinite Horizon Optimal Control Problems with Discounting. SIAM J Control Optim 48(4):2480-2512, 2009.

[16] D. Hernández-Hernández, O. Hernández-Lerma and M. Taksar. The linear programming approach to deterministic optimal control problems. Applicationes Mathematicae, 24(1):17–33, 1996.

[17] A. Jameson and E. Kreindler. Inverse problem of linear optimal control. SIAM Journal on Control, 11(1):1–19, 1973.

[18] R. E. Kalman. When is a linear control system optimal? Journal of Basic Engineering, 86(1):51–60, 1964.

[19] A. Keshavarz, Y. Wang, and S. P. Boyd. Imputing a convex objective function. International Symposium on Intelligent Control, IEEE, 2011.

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

[21] J.B. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear optimal control via occupation measures and LMI relaxations. SIAM Journal on Control and Optimization, 47(4):1643–1666, 2008.

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

[23] C. McDiarmid. On the method of bounded differences. In Surveys in Combinatorics, 148–188. Cambridge University Press, 1989.

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

[25] P. Moylan and B. D. O. Anderson. Nonlinear regulator theory and an inverse optimal control problem. IEEE Transactions on Automatic Control, 18(5):460–465, 1973.

[26] F. Nori and R. Frezza. Linear optimal control problems and quadratic cost functions estimation. Mediterranean Conference on Control and Automation, 2004.

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

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

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

[30] A.S. Puydupin-Jamin, M. Johnson, and T. Bretl. A convex approach to inverse optimal control and its application to modeling human locomotion. International Conference on Robotics and Automation, IEEE, 2012.

[31] N.D. Ratliff, J.A. Bagnell, and M.A. Zinkevich. Maximum margin planning. International Conference on Machine Learning, ACM, 2006.

[32] F. Thau. On the inverse optimum control problem for a class of nonlinear autonomous systems. IEEE Transactions on Automatic Control, 12(6):674–681, 1967.

[33] L. Vandenberghe and S. P. Boyd. Semidefinite programming. SIAM Review, 38(1):49–95, 1996.

[34] V.N. Vapnik. An overview of statistical learning theory. IEEE Transactions on Neural Networks, 10(5):988–999, 1999.

[35] R. Vinter and R. Lewis. The equivalence of strong and weak formulations for certain problems in optimal control. SIAM J Control Optim 16(4):546-570, 1978.

[36] R. Vinter. Convex duality and nonlinear optimal control. SIAM Journal on Control and Optimization, 31(2):518–538, 1993.

[37] H. Whitney. Geometric integration theory. Princeton Univ. Press, 1957.