Inverse Optimal Control for Differentially Flat Systems with Application to Locomotion Modeling

Navid Aghasadeghi and Timothy Bretl

Abstract— **Inverse optimal control is the problem of computing a cost function with respect to which observed trajectories of a given dynamic system are optimal. In this paper, we present a new formulation of this problem for the case where the dynamic system is differentially flat. We show that a solution is easy to obtain in this case, in fact reducing to finite-dimensional linear least-squares minimization. We also show how to make this solution robust to model perturbation, sampled data, and measurement noise, as well as provide a recursive implementation for online learning. Finally, we apply our new formulation of inverse optimal control to model human locomotion during stair ascent. Given sparse observations of human walkers, our model predicts joint angle trajectories for novel stair heights that compare well to motion capture data ($R^2 = 0.97, RMSE = 1.95$ degrees). These exemplar trajectories are the basis for an automated method of tuning controller parameters for lower-limb prosthetic devices that extends to locomotion modes other than level ground walking.**

I. Introduction

Inverse optimal control (IOC) is the problem of computing a cost function with respect to which observed trajectories of a given dynamic system are optimal. The cost function is then a concise model of "expert" behavior that can be used as the basis for imitation learning [1]. Applications of this approach include learning aerobatic helicopter maneuvers [2], learning skills like swinging a forehand with a humanoid robot [3], complex manipulation on a robot arm [4], [5], and human gesture learning [6].

Formulations of inverse optimal control vary depending on the class of dynamic system under consideration. The works [7], [8] develop IOC approaches for systems modeled as Markov decision processes (MDP). Moreover, IOC approaches for non-linear continuous-time systems have been developed for the deterministic [9] and stochastic [10] cases. Many of these formulations lead to methods of solution that are computationally complex, so recent work has focused on taking advantage of special structure of the systems such as linearly-solvable MDPs [11] .

In this paper, we introduce an approach to inverse optimal control for the class of differentially flat systems [12], [13]. For such systems, the deterministic non-linear optimal control problem can be equivalently written as an *unconstrained* calculus of variations problem. We show that this equivalence has five important consequences for the corresponding inverse optimal control problem. First, we demonstrate that the solution of the inverse optimal control problem becomes "easy", in fact reducing to finite-dimensional linear least-squares minimization. Second, we precisely derive conditions for unique estimation of the cost parameters given a set of demonstrations and basis functions for the cost. Third, we show that the problem of IOC for differentially flat systems does not explicitly depend on the dynamic equations, thus leading to a robust estimation in the presence of model perturbations. In this sense our algorithm is similar to model-free IOC algorithms [10], [14]. Fourth, we demonstrate how standard filtering techniques can be applied to the problem of IOC, to obtain better solutions in the presence of sampled data and measurement noise. Fifth, a recursive version of the IOC solution can be formulated for differentially flat systems, making this approach suitable for online robot learning applications.

An important motivation for developing the proposed IOC algorithm is to learn controllers for lower-limb prosthetic devices. A common strategy for control of these devices proceeds by breaking a gait cycle into four phases, and by applying a proportional derivative (PD) feedback policy within each phase [15]. A challenge in the application of this controller is that it requires the choice of many parameters, in particular 12 parameters for the knee joint. Currently clinicians choose these parameters by trial and error for each patient as noted in [15], [16]. This tuning process takes four hours for each locomotion mode, for e.g. level ground walking or stair/ramp ascent. In fact different controllers need to be tuned to provide locomotion for different inclinations of stairs and ramps [16].

In [17] we proposed a framework for automatically learning prosthetic controller parameters customized to an individual amputee for level ground walking. The proposed algorithm made use of a set of exemplar outputs corresponding to level walking, which have been shown to be invariant across subjects. These invariant outputs correspond to outputs of a five-link fully-actuated bipedal model, which is differentially flat [18]. We are interested in extending this method to learning prosthetic controllers for other locomotion modes such as ascending stairs and ramps of different inclinations. However, a major challenge is that we need exemplar locomotion outputs corresponding to any desired inclination of stair or ramp to use in the framework.

The proposed method of inverse optimal control exactly overcomes this challenge. By applying this method to the outputs of a five-link differentially flat biped, we can produce exemplar locomotion trajectories for stairs and ramps of

N. Aghasadeghi is with the Department of Electrical and Computer Engineering and T. Bretl is with the Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA {aghasad1,tbretl}@illinois.edu

different inclinations. These exemplar trajectories can then immediately be used in the prosthetic learning framework proposed in [17] to find prosthetic controller parameters for different modes of locomotion. The proposed method of IOC based on differential flatness is well-suited for this application. Due to the efficiency and scalability of the algorithm, we can learn cost functions composed of a large number of features. The learned cost function is also expected to work well for novel circumstances, due to the theoretical guarantees on estimation of unique cost parameters. Furthermore, the model-free nature of the algorithm is desirable since we often have data corresponding to multiple subjects, where the physical characteristics of the subjects, such as the mass of their body segments are not exactly known. Using the proposed method, we do not need to model these unknown and complex physical dynamics of the subjects. In this paper, we will describe how using sparse observations of stair ascent we can learn an IOC model. Subsequently, we demonstrate the power of the model in predicting locomotion outputs for a novel stair height and verify the predictions by comparing to motion captured data.

The remainder of the paper proceeds as follows. In section II we discuss an existing formulation of IOC for deterministic nonlinear systems. Theory of IOC for differentially flat systems is introduced in section III and is demonstrated on a simulated system in section IV. Subsequently, we apply the IOC method to model human stair ascent in section V and we conclude in section VI.

II. Existing Inverse Optimal Control Formulation and Solution

Consider the following deterministic non-linear continuous-time optimal control problem

$$\begin{aligned} \min_{x,u} \quad & \int_{t_0}^{t_f} \beta^{*T} \phi(x,u) dt \qquad (1) \\ \text{subject to} \quad & \dot{x} = f(x,u), \\ & x(t_0) = a, \\ & x(t_f) = b, \qquad (2) \end{aligned}$$

where $x \in \mathbb{R}^n$ represents the state of the system, $u \in \mathbb{R}^m$ represents the inputs, $f(x,u) : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ represents the system dynamics and $t_0, t_f \in \mathbb{R}^+$ represent the initial and final time. The known basis functions $\phi(x,u) : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^p$ together with the cost parameters $\beta^* \in \mathbb{R}^p$ represent a cost function to be minimized. Note that for simplicity of presentation, additional state-input constraints on the optimization (1) have not been included. However, both the existing method of IOC, and the proposed method are suitable for handling constraints, as we describe later.

A. Existing IOC Formulation

The problem of inverse optimal control for deterministic nonlinear systems, as noted in [9], [19], is often posed as finding the unknown cost parameters β^* given observations of the outputs $y = h(x, u, ..., u^{(j)}) \in \mathbb{R}^q$ for a finite j, and given known system dynamics $\dot{x} = f(x,u)$ and known basis functions $\phi(x,u)$.

The existing formulation of this approach discussed in [9], defines the following optimization to solve for the cost parameter

$$\arg\min_{\beta} \int_{t_0}^{t_f} ||y(t) - y(t,\beta)||^2 dt, \qquad (3)$$

where the norm is often chosen to be an L_2-norm. Here $y(t)$ is the observed output, and $y(t,\beta)$ is the solution to the following optimal control problem

$$\begin{aligned} \min_{x,u} \quad & \int_{t_0}^{t_f} \beta^T \phi(x,u) dt \qquad (4) \\ \text{subject to} \quad & \dot{x} = f(x,u), \\ & x(t_0) = a, x(t_f) = b, \\ & y(t,\beta) = h(x,u,...,u^{(j)}). \end{aligned}$$

B. Existing Solution Approach

An existing approach [9] to solving the optimization (3)-(4) proceeds by iteratively solving the optimal control problem (4) and updating the estimate of β. Thus in applications where solving the optimal control problem is computationally expensive, this iterative IOC approach is faced with a computational bottleneck. Other approaches such as [19], solve (3)-(4) by discretizing time and by forming a nonlinear program (NLP).

III. Inverse Optimal Control for Differentially Flat Systems

In this section, we introduce an IOC approach for the class of *differentially flat* systems. We first discuss how by focusing on the class of differentially flat systems, we can represent the optimal control problem (1) as an *equivalent* unconstrained calculus of variation problem. Subsequently for this class of systems, we introduce an alternative formulation of the IOC problem that is based on minimizing the extent to which first-order necessary conditions of optimality are violated.

A system is differentially flat [12], [13] if all the states x and inputs u can be obtained from the outputs without integration. More precisely, a system is differentially flat if for outputs $y = h(x, u, ..., u^{(j)}) \in \mathbb{R}^m$ there exists functions γ_x and γ_u such that

$$x = \gamma_x(y, \dot{y}, ..., y^{(k)}) \qquad (5)$$

$$u = \gamma_u(y, \dot{y}, ..., y^{(k)}) \qquad (6)$$

for some finite k-th order time derivative of the outputs denoted by $y^{(k)}$. Note that for a system to be differentially flat, the number of outputs q should equal the number of inputs m, i.e. the number of equations by which the ordinary differential equation (ODE) $\dot{x} = f(x,u)$ is underdetermined. In this paper, we will use the terms differentially flat and flat interchangeably. Some examples of flat systems include fully actuated Lagrangian systems such as a fully actuated five-link biped, the car with N trailers, and a kinematic chain [18].

For such systems, the optimal control problem (1) is equivalent to the following calculus of variation problem on the flat outputs as discussed in [20]–[22]

$$\min_{y,\ldots,y^{(k)}} \quad \int_{t_0}^{t_f} \beta^{*T}\phi(\gamma_x(y,\ldots,y^{(k)}),\gamma_u(y,\ldots,y^{(k)}))dt \tag{7}$$

$$\text{subject to} \quad \gamma_x(y(t_0),\ldots,y^{(k)}(t_0)) = a$$
$$\gamma_x(y(t_f),\ldots,y^{(k)}(t_f)) = b.$$

Now the problem of inverse optimal control for flat systems is to find the unknown parameters β^* given the observed flat outputs y and known basis function $\phi(\gamma_x(Y),\gamma_u(Y))$. Here $Y = [1, y, \dot{y}, ..., y^{(k)}]^T$ denotes the outputs and their higher order derivatives (we have also added the constant 1 to specify that the functions γ_x and γ_u can have constant terms independent of the outputs). Note that since the optimization (7) is no longer constrained by the dynamics, the corresponding IOC problem does not explicitly depend on the dynamic equations and only requires observations of the flat outputs.

A. Proposed IOC Formulation

In this section, we describe an alternative formulation of the IOC problem based on our previous work [23], [24], for the case where the system is differentially flat. We proceed by writing the Euler-Lagrange equations corresponding to the first order necessary conditions of optimality of (7) for an unknown parameter β as follows

$$\beta^T\sum_{\ell=0}^{k}(-1)^\ell\frac{d^\ell}{dt^\ell}\frac{\partial}{\partial y_1^{(\ell)}}\phi\Big(\gamma_x(Y),\gamma_u(Y)\Big) = 0,$$
$$\vdots$$
$$\beta^T\sum_{\ell=0}^{k}(-1)^\ell\frac{d^\ell}{dt^\ell}\frac{\partial}{\partial y_m^{(\ell)}}\phi\Big(\gamma_x(Y),\gamma_u(Y)\Big) = 0,$$

(where $y^{(0)} = y$ and $\frac{d^0}{dt^0}y = y$). We can write the necessary conditions of optimality succinctly as

$$\beta^T\sum_{\ell=0}^{k}(-1)^\ell\frac{d^\ell}{dt^\ell}\Phi_{y^{(\ell)}}\Big(\gamma_x(Y),\gamma_u(Y)\Big) = \mathbf{0} \tag{8}$$

where $\Phi_z(\gamma_x(Y),\gamma_u(Y))$ denotes the Jacobian of $\phi(\gamma_x(Y),\gamma_u(Y))$ with respect to the vector z, and $y^{(j)} = [y_1^{(j)}, y_2^{(j)}, ..., y_m^{(j)}]^T$ denotes the j-th derivative of the vector of outputs y. We further define the residual function $r(\beta, Y(t))$ as

$$r(\beta,Y(t)) = \beta^T\sum_{\ell=0}^{k}(-1)^\ell\frac{d^\ell}{dt^\ell}\Phi_{y^{(\ell)}}\Big(\gamma_x(Y),\gamma_u(Y)\Big)$$

To solve for the unknown parameters β, we formulate the IOC problem as minimizing the extent to which the first order necessary conditions of optimality are violated, i.e. by minimizing the L_2 norm of the residual as follows

$$\hat{\beta} = \arg\min_{\beta} \quad \int_{t_0}^{t_f}\Big|\Big|r(\beta,Y(t))\Big|\Big|^2 dt \tag{9}$$
$$\text{subject to} \quad \beta_1 = 1.$$

This IOC formulation can equivalently be seen as a parameter estimation problem for the ODE equation (8) given the observations Y. Different methods of ODE parameter estimation have been developed using maximum likelihood estimation [25] and least squares optimization [26]. The optimization (9) corresponds to a least squares method for ODE parameter estimation [26]. Also, note that since this formulation relies on the first-order necessary conditions of optimality, our approach can also be applied to observed *locally* optimal trajectories, similar to [27], while also ensuring unique estimation of the cost parameters discussed in section III-C.

Note that the constraint on the parameter vector β (here on the first element of the vector β) is required to ensure that we find a non-trivial solution, since otherwise the all zero solution will minimize the objective. Also note that the optimization (7) has the same solution if we scale the cost by a positive number. Therefore, IOC in principle can only find β^* up to a scaling. The inclusion of the constraint thus ensures the possibility of having a unique solution. The choice of the constraint $\beta_1 = 1$ requires domain knowledge. However, the only domain knowledge necessary beforehand is that $\beta_1 \neq 0$, and the exact value of β_1 need not be known, again because cost functions scaled by a positive number lead to the same solution.

We should also note that additional constraints on the flat outputs can be added to optimization (7). Constraints that impose additional necessary conditions of optimality as a function of the unknown cost parameters can be incorporated in the residual minimization. Otherwise if they are not functions of the cost parameters, they will not change the IOC optimization.

B. Proposed IOC Solution Approach based on Linear Least Squares

In this section, we derive the least squares solution to optimization (9). First we define $A^T(Y(t)) = -\Big[\sum_{\ell=0}^{k}(-1)^\ell\frac{d^\ell}{dt^\ell}\Phi_{y^{(\ell)}}\Big(\gamma_x(Y),\gamma_u(Y)\Big)\Big]_{2:p}$, and $b^T(Y(t)) = \Big[\sum_{\ell=0}^{k}(-1)^\ell\frac{d^\ell}{dt^\ell}\Phi_{y^{(\ell)}}\Big(\gamma_x(Y),\gamma_u(Y)\Big)\Big]_1$. Here the notation $\Big[X\Big]_1$ denotes the first row of the matrix X and $\Big[X\Big]_{2:p}$ denotes the second through p-th rows of the matrix X. Furthermore, we write $\beta^T = [1, \tilde{\beta}^T]$. The optimization (9) can equivalently be written as follows when we replace the constraint $\beta_1 = 1$ in the objective

$$\hat{\beta} = \arg\min_{\tilde{\beta}} \int_{t_0}^{t_f} ||b(Y(t)) - A(Y(t))\tilde{\beta}||^2 dt \tag{10}$$

The minimization can be written as

$$\min_{\tilde{\beta}} \int_{t_0}^{t_f} (b^T b - 2\tilde{\beta}^T A^T b + \tilde{\beta}^T A^T A \tilde{\beta}) dt.$$

Differentiating the convex objective with respect to the constant $\tilde{\beta}$ and equating to zero yields the equation

$$\Big(\int_{t_0}^{t_f} A^T(Y(t)) A(Y(t))\ dt \Big) \hat{\beta} = \Big(\int_{t_0}^{t_f} A^T(Y(t)) b(Y(t))\ dt \Big). \tag{11}$$

Equation (11), which is the continuous-time analog to the discrete-time linear least squares solution, thus solves the IOC problem in closed form. Note that standard Gaussian elimination or LU decomposition techniques can be used to find numerically stable solutions to the equation (11). Also note that if $A^T A$ is not bounded, we can get a bounded signal by normalization [28]. Thus from now on without loss of generality we will always assume $A^T A$ is bounded.

C. Properties of The Solution

To analyze the uniqueness of the solution, note that if the matrix $\int_{t_0}^{t_f} A^T A dt$ is full rank, a unique estimate $\hat{\beta}$ is obtained. Also note that the true cost parameter β^* satisfies the necessary conditions of optimality (8), and therefore guarantees the existence of a solution to the optimization (10). Therefore when the matrix $\int_{t_0}^{t_f} A^T A dt$ is full rank we can conclude that the unique estimate equals the true cost parameter, i.e. $\hat{\beta} = \beta^*$. In practice, one can check the condition number of the matrix instead of the rank to evaluate this condition.

D. Solution in the Presence of Noise and Sampled Data

Next, we will discuss two methods for computing an estimate of the cost parameters when the observed outputs are sampled and noisy. Both methods can lead to consistent estimation of the cost parameter, and the choice of the method depends on the application of interest. The first method will be based on spline-fitting the observed outputs, and the second method will make use of output filtering.

1) Solution via Spline-Fitting: Consider observing sampled noisy outputs. We model these observations as $z(k) \in \mathbb{R}^p$ where $k = 1, ..., N$ as follows

$$z(k) = y(t_k) + \epsilon_k \quad \text{where } t_0 \le t_1 \le ... \le t_N \le t_f \quad \text{and } \epsilon_k \sim N(0, \Sigma_p), \Sigma_p > 0. \tag{12}$$

The problem of finding an estimate of the cost parameter β in (8) can be considered as a parameter estimation problem for ODEs. An approach to parameter estimation using sampled noisy observations [26] proceeds in two steps

1) Find a consistent nonparametric estimate of $y(t)$, which we denote by $\tilde{y}(t)$ and its derivatives $\tilde{y}^{(n)}(t)$ using observations $z(k)$
2) Use the continuous-time estimates of the outputs to obtain an estimate of the unknown parameters

The problem of finding consistent nonparametric estimates has been largely studied and it is shown that such estimates can be constructed using splines and polynomial regression approaches [29], among others. Moreover, consistent estimates of the derivatives of $y(t)$ can also be obtained by differentiating $\tilde{y}(t)$ as described in [30].

Using the described two-step approach, it is shown [26] that we can find consistent estimates of the ODE parameters, which correspond to cost parameters in our work.

2) Solution via Output Filtering: In solving the IOC problem via optimization (10) we make use of higher order derivatives (HOD) of the observed outputs $y(t)$. Since we often only observe $y(t)$ and not its derivatives, we have to rely on numerical differentiation methods to estimate these HODs. This approach was used to do IOC via spline fitting.

In this section we develop an alternative method of IOC based on filtering, which does not require computing HODs. Note that the matrix $A(Y(t))$ and vector $b(Y(t))$ each have entries that are functions of $[y, \dot{y}, ..., y^{(r)}]$ up to some finite derivative of order r. We proceed by filtering the entries of $A(Y(t))$ and $b(Y(t))$ using an r-th order stable filter

$$\frac{1}{\Lambda(s)} = \frac{1}{s^r + \lambda_{r-1} s^{r-1} + \cdots + \lambda_0}. \tag{13}$$

We apply this filter to both sides of the necessary conditions of optimality (8) as follows

$$\frac{1}{\Lambda(s)} \mathcal{L}\{b(Y(t))\} = \frac{b_{\mathcal{L}}(s)}{\Lambda(s)} Y_{\mathcal{L}}(s)$$
$$\frac{1}{\Lambda(s)} \mathcal{L}\{A(Y(t))\} \tilde{\beta} = \frac{A_{\mathcal{L}}(s)}{\Lambda(s)} Y_{\mathcal{L}}(s) \tilde{\beta}$$

where $\mathcal{L}$ represents the Laplace operator, and $A_{\mathcal{L}}(s)$ and $b_{\mathcal{L}}(s)$ denote matrices of same size as $A(Y(t))$ and $b(Y(t))$ respectively. The entries of $A_{\mathcal{L}}(s)$ and $b_{\mathcal{L}}(s)$ are polynomial functions of s, and satisfy equations $A_{\mathcal{L}}(s) Y_{\mathcal{L}}(s) = \mathcal{L}\{A(Y(t))\}$ and $b_{\mathcal{L}}(s) Y_{\mathcal{L}}(s) = \mathcal{L}\{b(Y(t))\}$. Therefore, using this approach, we can treat every entry of $\frac{A_{\mathcal{L}}(s)}{\Lambda(s)}$ and $\frac{b_{\mathcal{L}}(s)}{\Lambda(s)}$ as a filter that is being applied to the observed outputs, and we eliminate the need for numerical computation of derivatives. Moreover, we can use the well-established theory of signal processing to design filters that can handle different kinds of measurement noise. By defining the filter outputs

$$\tilde{b}(t) = \mathcal{L}^{-1}\{\frac{b_{\mathcal{L}}(s)}{\Lambda(s)} Y_{\mathcal{L}}(s)\}$$
$$\tilde{A}(t) = \mathcal{L}^{-1}\{\frac{A_{\mathcal{L}}(s)}{\Lambda(s)} Y_{\mathcal{L}}(s)\},$$

we can now write the following IOC optimization, which can be solved as in (10)

$$\hat{\beta} = \arg\min_{\tilde{\beta}} \int_{t_0}^{t_f} ||\tilde{b}(t) - \tilde{A}(t)\tilde{\beta}||^2 dt. \tag{14}$$

To apply this method to sampled data $z(k)$ which are generated according to (12), with the assumption that $t_i - t_{i-1} = \Delta t, \forall i$, we can make use of the bilinear transformation [31] $s = \frac{2}{\Delta t}\frac{z-1}{z+1}$ to convert the continuous-time filter to a discrete-time filter.

E. Recursive Solution

Real-time learning is an important topic in robotics, since robots often encounter new situations and need to be equipped with the ability to self-improve [32]. In this section, we develop a recursive method of IOC for flat systems. This method will allow us to update the estimate of the cost parameters upon receiving new observations without resolving the complete IOC problem.

As we described, to solve IOC given data $Y(s)$ from time t_0 to t (i.e. $t_0 \le s \le t$) we solve the following optimization

$$\hat{\beta}(t) = \arg\min_{\tilde{\beta}} \int_{t_0}^{t} ||b(Y(s)) - A(Y(s))\tilde{\beta}||^2 ds.$$

To develop a recursive version of the IOC, we propose minimizing the following objective with respect to $\tilde{\beta}$

$$J(\tilde{\beta}) = \int_{t_0}^{t} ||b(Y(s)) - A(Y(s))\tilde{\beta}||^2 ds + \frac{1}{2}(\tilde{\beta} - \beta_0)^T Q_0 (\tilde{\beta} - \beta_0),$$

where we have added an extra cost to penalize deviation from some initial guess β_0 using a positive definite matrix $Q_0 > 0$. Note that given multiple demonstrations, we can solve the recursive IOC for the current demonstration, and use the estimated cost parameter as an initial guess when given the subsequent demonstration.

The objective $J(\tilde{\beta})$ is convex in $\tilde{\beta}$ and we set the gradient to zero to find the following equation that solves for $\hat{\beta}$ up to time t

$$\hat{\beta}(t) = \Big(\int_{t_0}^{t} A^T A ds + Q_0\Big)^{-1}\Big(\int_{t_0}^{t} A^T b ds + Q_0\beta_0\Big).$$

We define the matrix $P(t)$ as

$$P(t) := \Big(\int_{t_0}^{t} A^T A ds + Q_0\Big)^{-1}.$$

Note that $P(t)$ is well defined because $A^T A \ge 0$ and $Q_0 > 0$. To compute the cost parameter estimate $\hat{\beta}(t)$ recursively, we obtain an expression for $\dot{P}$ as

$$PP^{-1} = I \Rightarrow 0 = \frac{d}{dt}(PP^{-1}) = \dot{P}P^{-1} + PA^T A$$
$$\Rightarrow \dot{P} = -PA^T AP,$$

and subsequently obtain a differential equation for $\hat{\beta}$ as

$$\begin{aligned}\dot{\hat{\beta}} &= \dot{P}\Big(\int_{t_0}^{t} A^T b ds + Q_0\beta_0\Big) + PA^T b \\ &= -PA^T AP\Big(\int_{t_0}^{t} A^T b ds + Q_0\beta_0\Big) + PA^T b \\ &= -PA^T A\hat{\beta} + PA^T b.\end{aligned}$$

Thus the recursive IOC solution can be obtained from

$$\begin{aligned}\dot{\hat{\beta}} &= -PA^T A\hat{\beta} + PA^T b \\ \dot{P} &= -PA^T AP.\end{aligned} \tag{15}$$

Using theorem 4.3.2 in [28] we can show that the recursive solution $\hat{\beta}(t)$ converges to a constant $\bar{\beta}$, and if the matrix $A(t)^T$ is persistently exciting, the solution converges to the true cost parameter β^*. We say $A(t)^T$ is persistently exciting if for some $\alpha_0, T_0 > 0$ we have

$$\int_{t}^{t+T_0} A^T(s)A(s)ds \ge \alpha_0 T_0 I, \forall t.$$

IV. Unicycle Example

In this section, we show how to apply the proposed IOC by considering a unicycle example. We consider the following optimal control problem

$$\begin{aligned}\arg\min_{x,u,v} &\int_{t_0}^{t_f} \frac{\beta_1}{2}(x_1(t) - x_1^*)^2 \\ &+ \frac{\beta_2}{2}(x_2(t) - x_2^*)^2 + \frac{\beta_3}{2}v^2(t)dt, \\ \text{subject to} \quad & \dot{x}(t) = \begin{bmatrix} v(t)\cos(x_3(t)) \\ v(t)\sin(x_3(t)) \\ u(t)\end{bmatrix} \\ & x_1(0) = a_1, x_2(0) = a_2 \\ & x_1(1) = b_1, x_2(1) = b_2.\end{aligned} \tag{16}$$

Here $x(t) \in \mathbb{R}^3$ denotes the state of the unicycle at time t, and $u(t), v(t) \in \mathbb{R}$ denote the control inputs at time t. The unknown cost parameters consist of $\{\beta_1, \beta_2, \beta_3, x_1^*, x_2^*\}$. A unicycle is differentially flat [13] with flat outputs $y_1(t) = x_1(t)$ and $y_2(t) = x_2(t)$. Thus, the optimal control problem (16) can equivalently be written as

$$\begin{aligned}&\arg\min_{y_1,y_2} \int_{t_0}^{t_f} \phi_U(y_1(t), \dot{y}_1(t), y_2(t), \dot{y}_2(t))dt = \\ &\arg\min_{y_1,y_2} \int_{t_0}^{t_f} \Big[\frac{\beta_1}{2}(y_1(t) - x_1^*)^2 + \frac{\beta_2}{2}(y_2(t) - x_2^*)^2 \\ &\qquad + \frac{\beta_3}{2}(\dot{y}_1^2(t) + \dot{y}_2^2(t))\Big]dt \\ &\text{subject to} \quad y_1(0) = a_1, y_2(0) = a_2 \\ &\qquad\qquad\quad y_1(1) = b_1, y_2(1) = b_2.\end{aligned} \tag{17}$$

To find the cost parameters, we write the first order necessary conditions of optimality for (17) as follows

$$\begin{aligned}&\beta_1 y_1(t) - \beta_1 x_1^* - \beta_3 \ddot{y}_1(t) = 0, \\ &\beta_2 y_2(t) - \beta_2 x_2^* - \beta_3 \ddot{y}_2(t) = 0.\end{aligned} \tag{18}$$

We redefine the unknown parameters as $\tilde{\beta}_1 = \beta_1, \tilde{\beta}_2 = \beta_2, \tilde{\beta}_4 = -\beta_1 x_1^*$ and $\tilde{\beta}_5 = -\beta_2 x_2^*$ and we fix $\beta_3 = 1$. Now the necessary conditions of optimality are linear in the unknowns $\tilde{\beta}_i$ (although the original objective was nonlinear in the parameters) and we can solve

$$\Big(\int_0^1 \begin{bmatrix} y_1(t) & 0 \\ 0 & y_2(t) \\ 1 & 0 \\ 0 & 1\end{bmatrix}\begin{bmatrix} y_1(t) & 0 & 1 & 0 \\ 0 & y_2(t) & 0 & 1\end{bmatrix} dt\Big)\hat{\beta} =$$
$$\Big(\int_0^1 \begin{bmatrix} y_1(t) & 0 \\ 0 & y_2(t) \\ 1 & 0 \\ 0 & 1\end{bmatrix}\begin{bmatrix}\ddot{y}_1(t) \\ \ddot{y}_2(t)\end{bmatrix} dt\Big). \tag{19}$$

Now we consider the case that we only observe sampled noisy observations $z(k)$ and demonstrate the use of output filtering. The spline-fitting approach can be applied in a similar and straight-forward way as well. We proceed by applying the following stable third-order filter

$$\frac{1}{\Lambda(s)} = \frac{1}{(s+1)(s+2)(s+3)}$$

to the Laplace transform of equations (18) to get

$$\begin{aligned}\tilde{\beta}_1 \frac{1}{\Lambda(s)}(Y_1(s) - y_1(t_0)) + \tilde{\beta}_4 \frac{1}{s\Lambda(s)} = \\ \frac{s^2}{\Lambda(s)}Y_1(s) - \frac{s}{\Lambda(s)}y_1(t_0) - \frac{1}{\Lambda(s)}\dot{y}_1(t_0), \\ \tilde{\beta}_2 \frac{1}{\Lambda(s)}(Y_2(s) - y_2(t_0)) + \tilde{\beta}_5 \frac{1}{s\Lambda(s)} = \\ \frac{s^2}{\Lambda(s)}Y_2(s) - \frac{s}{\Lambda(s)}y_2(t_0) - \frac{1}{\Lambda(s)}\dot{y}_2(t_0).\end{aligned}$$

The four filters $\frac{1}{s\Lambda(s)}, \frac{1}{\Lambda(s)}, \frac{s}{\Lambda(s)}$ and $\frac{s^2}{\Lambda(s)}$ are converted to discrete-time filters using the discussed bilinear transformation. The IOC problem is then solved using the filtered outputs as described in Section III-D.2. A monte-carlo simulation was performed using 300 different boundary conditions and cost parameters that were uniformly randomly generated. Each experiment was repeated with different noise levels ranging from a signal-to-noise ratio (SNR) of 0 to 30. Fig. 1 shows that as we expect, the estimated cost parameter approaches its true value.

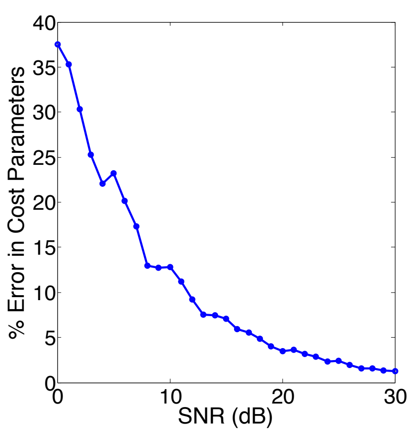

Fig. 1: IOC estimation error vs. SNR level

V. Application To Bipedal Locomotion and Prosthetic Controller Design

In this section, we will use the proposed method of IOC to obtain a model for human stair ascent, with application to learning controllers for lower-limb prosthetic devices. These devices are commonly controlled by impedance control. This strategy breaks the gait cycle into four phases and uses a PD feedback policy within each phase. Currently clinicians choose the PD gains, which consist of 12 parameters for the knee, through a time-consuming trial and error process. In [17] we proposed a framework to automate this tuning process for level walking. However, to extend this framework to other locomotion modes, we need a model that can generate exemplar locomotion trajectories, since such trajectories are not always available. We overcome this challenge by learning an IOC model capable of generating exemplar trajectories corresponding to any desired stair height.

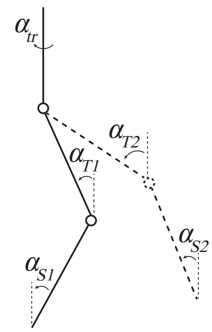

Fig. 2: Five Link Biped Model

We will make use of a five link fully actuated biped model shown in Fig. 2. Fully actuated mechanical systems are known to be differentially flat, with the joint angles as a candidate for flat outputs [33]. While this is a simple model, we will demonstrate that through the use of the proposed IOC algorithm it can learn and predict human locomotion trajectories. To derive a model of locomotion based on IOC, we make some design choices and will explain the motivation for these choices below.

A. Elevation Angles Instead of Joint Angles

A major challenge in learning lower-limb prosthetic controllers is that people with different physical characteristics (such as height and weight) walk with different joint trajectories. In [17] we overcome this problem by utilizing elevation angle trajectories as opposed to joint angles. These outputs have shown to be invariant across different subjects [34] and have been hypothesized to be controlled by the central nervous system [35]. The angles are shown in Fig. 2 and can be obtained from linear combination of joint angles, thus they also form a set of flat outputs. In Fig. 2 the subscripts T, S and tr denote thigh, shank and torso. Also stance is denoted by subscript 1 and swing by subscript 2.

B. Structure for the Cost Function

We model the cost for stair ascent as having a component that is independent of the stair height, and a component that is modulated by the stair height. We assume that the cost component corresponding to the thigh is directly modulated by the stair height, however, the cost component corresponding to the shank is independent of the stair height, and the shank angle is implicitly modulated through the thigh.

This cost structure is motivated by results in [36], [37] that propose the existence of kinetic and kinematic synergies between the shank and the thigh. Moreover, results in [38], [39] suggest that the thigh is modulated in a feedforward manner to produce different speeds and locomotion modes. We learn a cost function ϕ_L composed of the summation of the following parts for both stance and swing phases separately (We are omitting the stance and swing subscripts $i = 1, 2$ for notational simplicity.)

- $\phi_1(\alpha_S, \alpha_T)$: representing the component of the cost function capturing the relationship between the shank (α_S) and thigh (α_T) elevation angles.
- $\phi_2(\alpha_T, h)$: representing the component modulated by the stair height h.
- $\phi_3(\alpha_{tr})$: representing the cost for the torso.

Next we describe the different features that the functions ϕ_1, ϕ_2 and ϕ_3 are composed of weighted by cost parameters β_i.

1) Features Modeling Shank and Thigh Relationship: The cost component ϕ_1 equals the summation of the following features:

- $(\alpha_S - \beta_1)^2$: denotes the cost of regulation of the shank angle to a set point.
- $\beta_2 \dot{\alpha}_S^2$: denotes cost on angular velocity of the shank
- $\beta_3 \dot{\alpha}_T^2$: denotes cost on angular velocity of the thigh
- $\beta_4 \alpha_S \alpha_T + \beta_5 \alpha_S \alpha_T^2 + \beta_6 \alpha_S^2 \alpha_T$: denotes monomial basis functions modeling the relationship between the shank and the thigh.
- $\beta_7 \sin(\alpha_S) + \beta_8 \cos(\alpha_S) + \beta_9 \sin(\alpha_T) + \beta_{10} \cos(\alpha_T) + (\beta_{11} \sin(\alpha_S) + \beta_{12} \sin(\alpha_T))^2 + (\beta_{13} \cos(\alpha_S) + \beta_{14} \cos(\alpha_T))^2$: this feature corresponds to attractors for foot position, and can be derived by expanding cost on foot position regulation.

2) Features Modeling the Modulation of the Thigh: The thigh angle has an additional cost component $\phi_2(\alpha_T, h)$ which is a function of the stair height as defined below

- $(\alpha_T - \beta_{15}(h))^2$: denotes the cost of regulation of the thigh angle to a set point.

3) Features Modeling the Torso: We define the following features for the torso

- $(\alpha_{tr} - \beta_{16})^2$: cost of regulation of torso to a set point.
- $\beta_{17} \dot{\alpha}_{tr}^2$: denotes cost on angular velocity of the torso.
- $\beta_{18} \sin(\alpha_{tr}) + \beta_{19} \cos(\alpha_{tr})$: denotes sinusoidal basis functions.

Now given the discussed locomotion cost $\phi_L = \phi_1 + \phi_2 + \phi_3$ we consider calculus of variation problem

$$\min_{\alpha_S, \alpha_T, \alpha_{tr}} \int_{t_0}^{t_f} e^{-t} (\phi_1(\alpha_S, \alpha_T) + \phi_2(\alpha_T, m) + \phi_3(\alpha_{tr}, m)) dt,$$

where e^{-t} represents a discounting factor. We apply the proposed method of IOC via cubic splines to solve for the unknown β cost parameters. The parameter $\beta_{15}(h)$ for swing is then varied to generate stair ascent trajectories for different inclinations (for stance β_{15} is independet of h).

C. Results and Predicting of Novel Stair Height

We use motion-captured data collected in [40] to learn an IOC model. These trajectories correspond to average joint trajectories of ten unimpaired subjects, climbing stairs at minimum, medium and maximum heights. We used output trajectories corresponding to the minimum and maximum stair heights as observations. The proposed method of IOC is then applied to estimate the cost parameters given these two observations. To evaluate the performance of our estimated cost, we evaluate how well we can predict the observed trajectories in terms of R^2 and $RMSE$ as seen in Fig. 3. Moreover, the model predicted a novel stair height (medium), and demonstrated a good fit to motion capture data ($R^2 = 0.97$ and $RMSE = 1.95$ degrees.)

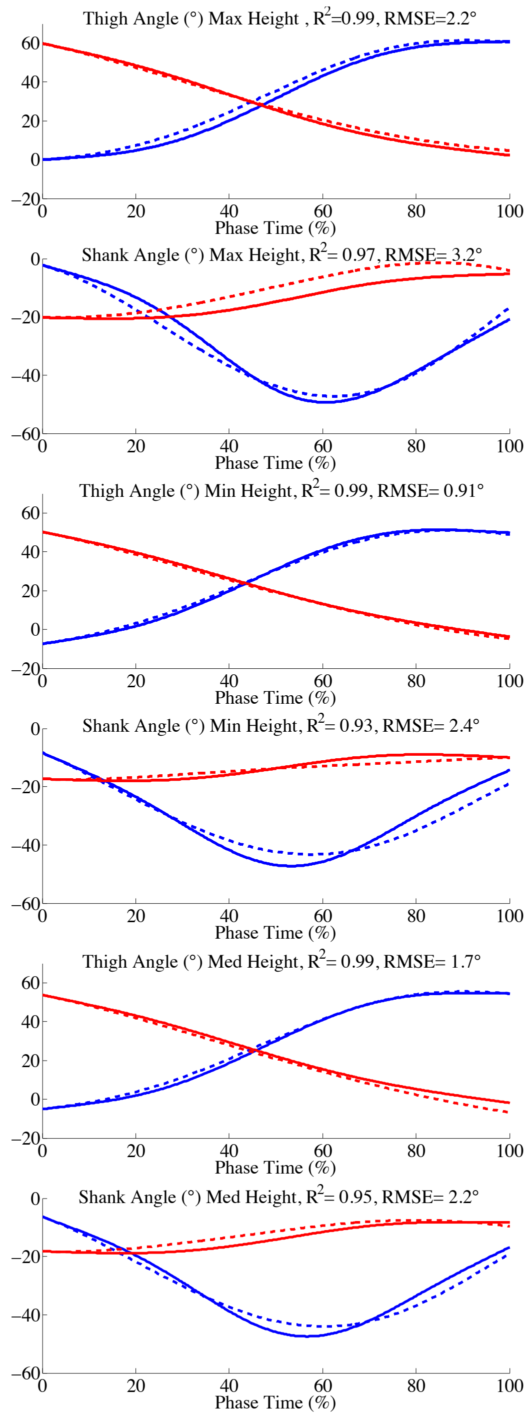

Fig. 3: Real (solid line) and Predicted (dashed line) stair ascent trajectories. Swing is shown in blue and stance in red. Quality of predictions is noted by R^2 and $RMSE$ on each plot. Data for max and min stair heights were used in learning, and medium height was a novel prediction by IOC.

VI. Conclusion

In this paper we introduced a new formulation of the problem of inverse optimal control for the case where the system is differentially flat. We showed that IOC for this class of systems has a number of desirable properties, including an efficient finite-dimensional linear least squares solution, robustness with respect to model perturbations and noise in sampled observations. We further demonstrated that our approach can learn a model of stair ascent using sparse data, and generate locomotion trajectories for novel stair heights. Thus this model enables the automated tuning of lower-limb prosthetic controllers for different locomotion modes. The scalability and robustness of this approach, along with its prediction power even with simple dynamic models, opens the door for many exciting robotic applications.

VII. Acknowledgments

The authors would like to thank Robert Riener and his research group for kindly sharing their experimental data. This work was supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under award #F31NS080618.

References

[1] S. Schaal, "Is imitation learning the route to humanoid robots?" *Trends in cognitive sciences*, vol. 3, no. 6, pp. 233–242, 1999.

[2] P. Abbeel, A. Coates, and A. Ng, "Autonomous helicopter aerobatics through apprenticeship learning," *International Journal of Robotics Research*, vol. 29, no. 13, pp. 1608–1639, 2010.

[3] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, vol. 2. IEEE, 2002, pp. 1398–1403.

[4] F. Stulp, E. Theodorou, M. Kalakrishnan, P. Pastor, L. Righetti, and S. Schaal, "Learning motion primitive goals for robust manipulation," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 325–331.

[5] E. Gribovskaya, S. M. Khansari-Zadeh, and A. Billard, "Learning non-linear multivariate dynamics of motion in robotic manipulators," *The International Journal of Robotics Research*, vol. 30, no. 1, pp. 80–117, 2011.

[6] S. Calinon, F. D'halluin, E. L. Sauser, D. G. Caldwell, and A. G. Billard, "Learning and reproduction of gestures by imitation," *Robotics & Automation Magazine, IEEE*, vol. 17, no. 2, pp. 44–54, 2010.

[7] P. Abbeel and A. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proceedings of the twenty-first international conference on Machine learning*. ACM, 2004, p. 1.

[8] B. Ziebart, A. Maas, J. Bagnell, and A. Dey, "Maximum entropy inverse reinforcement learning," in *Proc. AAAI*, 2008, pp. 1433–1438.

[9] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion—an inverse optimal control approach," *Autonomous Robots*, vol. 28, no. 3, pp. 369–383, 04 2010.

[10] N. Aghasadeghi and T. Bretl, "Maximum entropy inverse reinforcement learning in continuous state spaces with path integrals," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, 2011, pp. 1561–1566.

[11] K. Dvijotham and E. Todorov, "Inverse Optimal Control with Linearly-Solvable MDPs," in *Proceedings of the Interntional Conference on Machine Learning*. Citeseer, 2010.

[12] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, "Flatness and defect of non-linear systems: introductory theory and examples," *International journal of control*, vol. 61, no. 6, pp. 1327–1361, 1995.

[13] R. M. Murray, M. Rathinam, and W. Sluis, "Differential flatness of mechanical control systems: A catalog of prototype systems," in *ASME International Mechanical Engineering Congress and Exposition*. Citeseer, 1995.

[14] A. Boularias, J. Kober, and J. Peters, "Relative entropy inverse reinforcement learning," in *International Conference on Artificial Intelligence and Statistics*, 2011, pp. 182–189.

[15] F. Sup, A. Bohara, and M. Goldfarb, "Design and control of a powered transfemoral prosthesis," *The International journal of robotics research*, vol. 27, no. 2, pp. 263–273, 2008.

[16] F. Sup, H. A. Varol, and M. Goldfarb, "Upslope walking with a powered knee and ankle prosthesis: initial results with an amputee subject," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 19, no. 1, pp. 71–78, 2011.

[17] N. Aghasadeghi, H. Zhao, L. J. Hargrove, A. D. Ames, E. J. Perreault, and T. Bretl, "Learning impedance controller parameters for lower-limb prostheses," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on, accepted*, 2013.

[18] R. M. Murray, "Nonlinear control of mechanical systems: A lagrangian perspective," *Annual Reviews in Control*, vol. 21, pp. 31–42, 1997.

[19] K. Hatz, J. P. Schloder, and H. G. Bock, "Estimating parameters in optimal control problems," *SIAM Journal on Scientific Computing*, vol. 34, no. 3, pp. A1707–A1728, 2012.

[20] S. Agrawal and N. Faiz, "Optimization of a class of nonlinear dynamic systems: new efficient method without lagrange multipliers," *Journal of optimization theory and applications*, vol. 97, no. 1, pp. 11–28, 1998.

[21] R. Mahadevan, S. K. Agrawal, and F. J. Doyle, "Differential flatness based nonlinear predictive control of fed-batch bioreactors," *Control Engineering Practice*, vol. 9, no. 8, pp. 889–899, 2001.

[22] J. Oldenburg and W. Marquardt, "Flatness and higher order differential model representations in dynamic optimization," *Computers & chemical engineering*, vol. 26, no. 3, pp. 385–400, 2002.

[23] N. Aghasadeghi, A. Long, and T. Bretl, "Inverse optimal control for a hybrid dynamical system with impacts," in *Robotics and Automation (ICRA), 2012 IEEE/RSJ International Conference on*. IEEE, 2012.

[24] M. Johnson, N. Aghasadeghi, and T. Brelt, "Inverse optimal control for deterministic continuos-time non-linear systems," in *Decision and Control, 2013. Proceedings of the 52nd IEEE Conference on, accepted*. IEEE, 2013.

[25] J. Ramsay, G. Hooker, D. Campbell, and J. Cao, "Parameter estimation for differential equations: a generalized smoothing approach," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 5, pp. 741–796, 2007.

[26] N. Brunel, "Parameter estimation of odes via nonparametric estimators," *Electronic Journal of Statistics*, vol. 2, pp. 1242–1267, 2008.

[27] S. Levine and V. Koltun, "Continuous inverse optimal control with locally optimal examples," *arXiv preprint arXiv:1206.4617*, 2012.

[28] P. A. Ioannou and J. Sun, *Robust adaptive control*. Courier Dover Publications, 2012.

[29] A. Berlinet and C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in probability and statistics*. Kluwer Academic Boston, 2004.

[30] S. Zhou and D. Wolfe, "On derivative estimation in spline regression," *Statistica Sinica*, vol. 10, no. 1, pp. 93–108, 2000.

[31] A. V. Oppenheim, R. W. Schafer, J. R. Buck et al., *Discrete-time signal processing*. Prentice Hall Upper Saddle River, 1999, vol. 5.

[32] S. Schaal, C. G. Atkeson, and S. Vijayakumar, "Scalable techniques from nonparametric statistics for real time robot learning," *Applied Intelligence*, vol. 17, no. 1, pp. 49–60, 2002.

[33] P. Martin, R. M. Murray, P. Rouchon et al., "Flat systems," 1997.

[34] N. A. Borghese, L. Bianchi, and F. Lacquaniti, "Kinematic determinants of human locomotion." *The Journal of physiology*, vol. 494, no. Pt 3, pp. 863–879, 1996.

[35] F. Lacquaniti, R. Grasso, and M. Zago, "Motor patterns in walking," *Physiology*, vol. 14, no. 4, pp. 168–174, 1999.

[36] D. A. Winter, *Biomechanics and motor control of human gait: normal, elderly and pathological*, 1991.

[37] Y. P. Ivanenko, A. d'Avella, R. E. Poppele, and F. Lacquaniti, "On the origin of planar covariation of elevation angles during human locomotion," *Journal of neurophysiology*, vol. 99, no. 4, pp. 1890–1898, 2008.

[38] M. Daley, G. Felix, and A. Biewener, "Running stability is enhanced by a proximo-distal gradient in joint neuromechanical control," *Journal of Experimental Biology*, vol. 210, no. 3, pp. 383–394, 2007.

[39] M. MacKay-Lyons, "Central pattern generation of locomotion: a review of the evidence," *Physical Therapy*, vol. 82, no. 1, pp. 69–83, 2002.

[40] R. Riener, M. Rabuffetti, and C. Frigo, "Stair ascent and descent at different inclinations," *Gait & posture*, vol. 15, no. 1, pp. 32–44, 2002.