

Assessing the Quality of a Set of Basis Functions for Inverse Optimal Control via Projection onto Global Minimizers

Filip Bečanović^{1,4,*} Jared Miller^{2,*} Vincent Bonnet³ Kosta Jovanović⁴ Samer Mohammed¹

Abstract— Inverse optimization (Inverse optimal control) is the task of imputing a cost function such that given test points (trajectories) are (nearly) optimal with respect to the discovered cost. Prior methods in inverse optimization assume that the true cost is a convex combination of a set of convex basis functions and that this basis is consistent with the test points. However, the consistency assumption is not always justified, as in many applications the principles by which the data is generated are not well understood. This work proposes using the distance between a test point and the set of global optima generated by the convex combinations of the convex basis functions as a measurement for the expressive quality of the basis with respect to the test point. A large minimal distance invalidates the set of basis functions. The concept of a set of global optima is introduced and its properties are explored in unconstrained and constrained settings. Upper and lower bounds for the minimum distance in the convex quadratic setting are implemented by bi-level gradient descent and an enriched linear matrix inequality respectively. Extensions to this framework include max-representable basis functions, nonconvex basis functions (local minima), and applying polynomial optimization techniques.

I. INTRODUCTION

There is a consensus in the literature on biomechanics, human motor control, and robotics to say that numerous human motions can be modeled and predicted using optimal control approaches [1–6]. This is especially the case for repetitive motions or tasks that are subject to strict biomechanical constraints such as walking [7]. To determine the cost function(s) used by humans, the use of Inverse Optimal Control (IOC) methods has been proposed and extensively studied [1–6]. Let a general motion be represented as a vector x from \mathbb{R}^n , and in particular, let observations be denoted by y . The application of IOC requires at least one (nearly) optimal observation $y \in \mathbb{R}^n$, knowledge of the constraints related to the biomechanics of the task and

of the human body that define the set of feasible motions $X \subseteq \mathbb{R}^n$, and a finite-cardinality set of basis cost functions $\mathcal{F} = \{f_i(x)\}_{i=1}^m$. Under the assumption that the cost function underlying human motion is a convex combination of functions from the basis $f(x) = \sum_{i=1}^m \alpha_i f_i(x)$, the goal of IOC is to identify parameters α_i such that the observation(s) y is (are) optimal. The basis functions are chosen based on physiological observations and are often related to the mechanical energy or motion smoothness [1].

IOC is often solved in human motor control or robotics by using a computationally expensive bi-level optimization approach [1–3, 7]. The bi-level approach consists of two optimization processes, one nested inside the other, that explicitly aim to minimize over the parameters α the Root-Mean-Square-Error between optimal observation y and the result of the so-called Direct Optimal Control (DOC) problem that returns the optimal solution x^* given the cost function $f(x)$. This approach tends to be abandoned due to its long execution time, which can take up to several days, but has the merit of being able to deal with relatively noisy input data [8] since the explicit minimization of the distance between y and the result of the inner optimization problem is accounted for.

Recently, an approximate solution method for the IOC problem was proposed based on a relaxation of the Karush-Kuhn-Tucker (KKT) optimality conditions, allowing to look for solutions by solving a linear least-square KKT-residual minimization [9]. Even for high-dimensional non-linear models of the DOC, the IOC admits an almost instantaneous solution, but requires the observation y to be within, or very close, to the set of possible solutions of the DOC denoted \mathcal{G}^c later (or \mathcal{G} in the absence of constraints). This method has motivated a fair amount of work in robotics [10–13] and human-motion analysis [4–6]. Many of these works are grounded in the analysis of the properties of the technique itself [14–16]. The method is based on a relaxation of the exact feasibility problem of inverse optimization, formulated as Problem 5 (resp. 3), for a constrained (resp. unconstrained) direct optimization problem.

Though these approximate methods [9, 10, 17] have considerable computational advantages with respect to the bi-level approach, they are known to be sensitive to noise [15, 18] and are statistically inconsistent [8]. However, when dealing with human motion, inherent simplifications are necessary in the modeling of the complex musculoskeletal system which introduces modelling errors that add up with non negligible measurement errors. For example, joint angles, that are often used as optimal observation y for IOC,

* Denotes equal contribution.

¹F. Bečanović, S. Mohammed are with the Université Paris-Est Créteil, Créteil, France (filip.becanovic@u-pec.fr, samer.mohammed@u-pec.fr).

²J. Miller is with the Robust Systems Laboratory, ECE Department, Northeastern University, Boston, MA 02115 (miller.jare@northeastern.edu).

³V. Bonnet is with the Laboratory of Analysis and Architecture of Systems, Toulouse (vincent.bonnet@laas.fr).

⁴F. Bečanović, K. Jovanović are with the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia (kostaj@etf.rs).

F. Bečanović was supported by the scholarship *Bourse de Thèse en Coutelle* of the *Institut Français en Serbie*. This research was partially supported by the Science Fund of the Republic of Serbia, PROMIS, #6062528, ForNextCobot.

J. Miller was partially supported by NSF grants CNS-1646121, ECCS-1808381 and CNS-2038493, AFOSR grant FA9550-19-1-0005, and ONR grant N00014-21-1-2431. J. Miller was in part supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States and the AFOSR International Student Exchange Program.

can have several degrees of error at the hip level during ground walking [19].

Few studies [16] propose to evaluate how well a basis of cost functions \mathcal{F} can represent a given motion y . One way could be to perform repeated IOC-DOC cycles with different initialization for either the bi-level or the approximate approaches, and look for the smallest obtainable error. Another approach would be to use the residual of the approximate KKT approach [9]. Unfortunately, what constitutes a good value for the residual can wildly vary from problem to problem, and even within the same problem depending on the conditioning [16].

This work proposes to evaluate how well a basis of cost functions \mathcal{F} can represent a given motion y . To make this concept intuitive, the concept of the *set of global minima* \mathcal{G}^c (\mathcal{G}) is introduced. The problem of evaluating a basis of cost functions \mathcal{F} is then formulated as the minimum-distance projection from the observation y to the set of global minima \mathcal{G}^c (\mathcal{G}). The projection is equivalent to finding the global minima of the bi-level problem, which is in general non-convex admitting no reliable way of finding global solutions. However, one can bound this distance from above, and from below under some assumptions. While the distance between any point from \mathcal{G}^c to y yields an upper bound on the minimum distance, one should strive to make the bound tighter by picking at least a local minimum of the bi-level problem. An approach based on an enhanced LMI will be presented for bounding the distance from below. As medium-sized LMIs are readily solved, this approach applies to problems of moderate sizes.

The set of global minima \mathcal{G}^c (\mathcal{G}) is characterized for convex bases \mathcal{F} , both with constrained and unconstrained DOC models. The problem of checking whether there exists a cost function parametrization α which makes the observation y exactly optimal is formulated as a linear program. The projection problem, equivalent to the bi-level approach, is then re-formulated as a single-level optimization process using the classical KKT-transformation of bi-level problems [20]. This reformulation allows for the application of enhanced LMIs for lower bounding the distance, in cases where the DOC model is polynomial in cost and constraints.

Further considerations are restricted to unconstrained and constrained quadratic programs as DOC models. Meaning the members of the set of basis functions \mathcal{F} are assumed quadratic and the constraints are, if existent, assumed linear. The geometry of the corresponding sets \mathcal{G} and \mathcal{G}^c are investigated, as well as the mappings between optimal motions x , cost function parametrizations α , and the dual variables λ and μ when constraints are present, through the extended global minima sets $\hat{\mathcal{G}}$ and $\hat{\mathcal{G}}^c$.

In this context, the contributions of this work are,

- The formulation of the Projection onto Global Minimizers (ProjGM) problem 2,
- The characterization of global optima set geometry,
- The use of numerical algorithms to determine upper and lower bounds in the convex quadratic setting.

This paper has the following structure: Section II defines

notation and introduces the problems that will be addressed in this work. Section III describes the structure of unconstrained global-optima sets, with specific reference to cases where each f_j is a strongly convex quadratic. Section IV analyzes constrained global-optima sets, and similarly focuses on the Quadratic Program (QP) case where the objective functions are weakly convex. Section V acquires upper bounds for program 2 using local search and lower bounds through an Linear Matrix Inequality (LMI). Section V presents a numerical example of the proposed ProjGM approach. Section VII details extensions to the current method, including adding support for maximum-representable functions in the cost functions set \mathcal{F} .

II. PRELIMINARIES

A. Notations

The set of real numbers is \mathbb{R} , the n -dimensional Euclidean space is \mathbb{R}^n , and its nonnegative real orthant is \mathbb{R}_+^n . The vector of all zeros is 0 and of all ones is $\mathbf{1}$. The set of natural numbers (incl. 0) is \mathbb{N} , and its subset between 1 and N is $1..N$. The n -dimensional probability simplex is Δ^n . The set of $m \times n$ real-valued matrices is $\mathbb{R}^{m \times n}$. The set of $n \times n$ symmetric real matrices satisfying $Q = Q^T$ is \mathbb{S}^n . A symmetric matrix is Positive Semidefinite (PSD) ($Q \in \mathbb{S}_+^n$) if $\forall x \in \mathbb{R}^n : x^T Q x \geq 0$, and is Positive Definite (PD) ($Q \in \mathbb{S}_{++}^n$) if $\forall x \in \mathbb{R}^n, x \neq 0 : x^T Q x > 0$.

Let X and Y be a pair of spaces. A single valued function may be written as $f : X \rightarrow Y$, and a set valued function $F : 2^X \rightarrow 2^Y$ may be expressed as $F : X \rightrightarrows Y$. The projection $\pi^x : (x, y) \mapsto x$ operation applied to a set $\hat{\mathbb{J}} \subseteq X \times Y$ is,

$$\mathbb{J}(x) = \pi^x \hat{\mathbb{J}}(x, y) = \{x \mid \exists y \in Y : (x, y) \in \hat{\mathbb{J}}\}. \quad (1)$$

B. Problem Statement

A convex combination cost f_α of basis functions from $\mathcal{F} = \{f_i(x)\}_{i=1}^m$ given weights α may be expressed as,

$$f_\alpha(x) = \sum_{j=1}^m \alpha_j f_j(x) \quad \alpha \in \Delta^m \quad (2)$$

where Δ^m is the m -dimensional probability simplices $\Delta^m = \{\alpha \in \mathbb{R}^m \mid \alpha \geq 0, \sum_i \alpha_i = 1\}$. This paper attempts to answer the following two problems:

Problem 1 (Feasibility): Does there exist an $\alpha \in \Delta^m$ such that y is a global optimum?

$$\underset{\alpha \in \Delta^m}{\text{find}} \ y \in \underset{x \in X}{\text{argmin}} \ \sum_{j=1}^m \alpha_j f_j(x) \quad (3)$$

Problem 2 (ProjGM): What is the distance from y to the set of global optima?

$$p^* = \min_{x \in X} \|y - x\|_2^2 \quad (4a)$$

$$\exists \alpha \in \Delta^m \mid x \in \underset{x' \in X}{\text{argmin}} \ \sum_{j=1}^m \alpha_j f_j(x') \quad (4b)$$

If α solves the feasibility program 1, then the objective from problem 2 will necessarily be $p^* = 0$ with $x = y$. Both problems will be treated under the following assumption (the same setting as used in [9]),

Assumption 1 (Convexity): Each f_j is a C^1 convex function and X is a convex set.

Problem 1 under Assumption 1 may be decided by analyzing the feasibility of a simple n -dimensional Linear Program (LP) in α . Problem 2 is a bilevel optimization problem that assesses the quality of the basis \mathcal{F} from (2) in characterizing y as an optimal point. A large distance p^* from Program (4) indicates that the basis \mathcal{F} does not accurately describe y as a global minimum and therefore \mathcal{F} should be redesigned. Problem 2 is a generically nonconvex and nontrivial problem even under Assumption 1. Local search and trust-region methods may be used to find upper bounds for p^* , but exact computation of p^* is typically intractable. An LMI is introduced to obtain lower bounds for p^* , and this LMI relaxation is tight to p^* if the PSD matrix variable is rank-1.

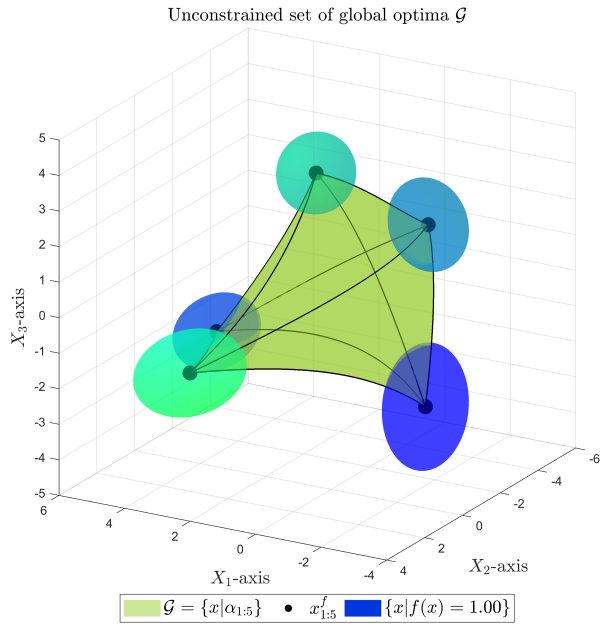


Fig. 1: A set of unconstrained global optima in 3D

III. GEOMETRY OF UNCONSTRAINED OPTIMAL SETS

This section formulates and defines properties of sets of global optima when $X = \mathbb{R}^n$. Due to repeated usage throughout this section, the symbol \mathcal{W} will denote $\mathbb{R}^n \times \Delta^m$.

A. Convex Case

The set of global optima and weights associated with a \mathcal{F} yielding mixed cost functions (2) given Assumption 1 is,

$$\hat{\mathcal{G}} = \{(x, \alpha) \in \mathcal{W} \mid \nabla_x f_\alpha(x) = 0\}. \quad (5)$$

Convexity of f_α implies that every local minimizer with $\nabla_x f_\alpha(x) = 0$ is also a global minimizer.

As condition $\nabla_x f_\alpha(y) = 0$ is linear in α , the feasibility problem 1 given $y \in \mathbb{R}^n$ may be posed as an LP in α ,

Problem 3 (Unconstrained Feasibility):

$$\text{find}_{\alpha \in \Delta^m} (y, \alpha) \in \hat{\mathcal{G}}. \quad (6)$$

The unconstrained ProjGM problem for the same y is,

Problem 4 (Unconstrained ProjGM):

$$p^* = \min_{(x, \alpha) \in \hat{\mathcal{G}}} \|y - x\|_2^2. \quad (7)$$

B. Quadratics Only

Further analysis in this section will require the following assumption,

Assumption 2 (Strictly Convex Quadratics): Functions in \mathcal{F} are strictly convex quadratics with,

$$f_j(x) = (x - x_j^f)^T Q_j (x - x_j^f) / 2 \quad \forall j = 1..m \quad (8a)$$

$$\nabla_x f_j(x) = Q_j (x - x_j^f) \quad \forall j = 1..m. \quad (8b)$$

where all elements in (Q_j, x_j^f) are bounded and $Q_j \in \mathbb{S}_{++}^n$ for all $j = 1..m$.

The optima-weight sets $\hat{\mathcal{G}}(x, \alpha)$ and $\mathcal{G}(x)$ under Assumption 2 is,

$$\hat{\mathcal{G}} = \{(x, \alpha) \in \mathcal{W} \mid \left(\sum_j \alpha_j Q_j\right) x = \sum_j \alpha_j Q_j x_j^f\} \quad (9)$$

$$\mathcal{G} = \pi^x \hat{\mathcal{G}}. \quad (10)$$

Remark 1: Letting $\{e_j\}_{j=1}^m$ be the standard basis vectors in \mathbb{R}^m , the points $\{(x_j^f, e_j)\}_{j=1}^m$ are all members of $\hat{\mathcal{G}}$.

Remark 2: All descriptor constraints in $\hat{\mathcal{G}}$ from (9) have polynomial degree at most two, and there is a bilinearity between (α, x) .

The mapping $\kappa : \Delta^m \rightarrow \mathcal{G}$ with $\kappa(\alpha) = x_\alpha^*$ in terms of finding an optimal x minimizing f_α is single-valued, and has an expression,

$$\kappa(\alpha) = x_\alpha^* = \left(\sum_j \alpha_j Q_j\right)^{-1} \left(\sum_j \alpha_j Q_j x_j^f\right). \quad (11)$$

Remark 3: Conversely, the mapping $\kappa^{-1} : \mathcal{G} \rightrightarrows \Delta^m$ is set-valued. Its values are the compact polytopic sets containing feasible points of program (6).

Theorem 3.1: The map $\kappa(\alpha)$ from (11) is a continuous surjection from Δ^m onto \mathcal{G} under Assumption 2.

Proof: Surjection holds by definition of $\hat{\mathcal{G}}$ in (9): x is only a member of $\hat{\mathcal{G}}$ if there exists an $\alpha' \in \Delta^m$ such that $x = x_{\alpha'}^*$. Continuity of x_α^* is based on continuity of the matrix inverse A^{-1} for all nonsingular matrices $A \in \mathbb{R}^{n \times n}$. ■

Definition 3.1 (Compact): A set $X \in \mathbb{R}^n$ is *compact* if it is closed and bounded (Heine-Borel). A consequence is that X is compact if there exists a finite $R > 0$ such that X is a subset of the ball with radius R : $\{x \mid \|x\|_2 < R\}$ [21].

Theorem 3.2: Under Assumption 2, the set $\hat{\mathcal{G}}$ is compact.

Proof: There exists a finite quantity $R > 0$ that satisfies $\max_j \|x_j^f\|_2 < R$ due to boundedness of x^f . It is implied that $\max_{\alpha \in \Delta^m} \|\sum_{j=1}^m \alpha_j x_j^f\|_2 < R$ by convexity of the norm $\|\cdot\|_2$. Let Λ be the solution to,

$$\Lambda = \min_{\alpha \in \Delta^m} \lambda_{\min} \left(\sum_j \alpha_j Q_j\right). \quad (12)$$

It holds that $\Lambda > 0$ because PD matrices form a convex (non-pointed) cone. The maximum norm of any global-optimal point in \mathcal{G} is bounded above by,

$$\|\kappa(\alpha)\|_2 = \left\| \left(\sum_j \alpha_j Q_j\right)^{-1} \left(\sum_j \alpha_j Q_j x_j^f\right) \right\|_2 \quad (13a)$$

$$\leq (1/\Lambda) \left\| \left(\sum_j \alpha_j Q_j x_j^f\right) \right\|_2 \leq R/\Lambda < \infty. \quad (13b)$$

The compact set $\{(x, \alpha) \in \mathcal{W} \mid \|x\|_2 \leq R/\Lambda\}$ is a superset of $\hat{\mathcal{G}}$ from (9), which proves that $\hat{\mathcal{G}}$ is compact. ■

Definition 3.2 (Path-Connected): A set X is path-connected if for every two points $x^0, x^1 \in X$ there exists a continuous path (curve) $\omega : [0, 1] \rightarrow X$ with $\omega(0) = x^0$, $\omega(1) = x^1$ such that $\omega(t) \in X \forall t \in [0, 1]$ [21].

Theorem 3.3: Under Assumption 2, the set \mathcal{G} is path-connected.

Proof: Let x^0, x^1 be a pair of distinct points in \mathcal{G} . Choose $\alpha^0 \in \kappa^{-1}(x^0)$ and $\alpha^1 \in \kappa^{-1}(x^1)$ as weights generating the optimal points x^0, x^1 . A path $\omega : [0, 1] \rightarrow \Delta^m$ may be drawn between the points by $\omega(t) : \alpha^0 t + \alpha^1(1-t)$. The path $(\omega(t), \kappa(\omega(t)))$ remains inside $\hat{\mathcal{G}}$ for all $t \in [0, 1]$ by continuity of κ from Theorem 3.1. This containment holds for all pairs (x^0, x^1) , so \mathcal{G} is path-connected. ■

Remark 4: No conclusions can be drawn in this manner about path-connectedness of $\hat{\mathcal{G}}$.

Figure 1 depicts the set of unconstrained global minima \mathcal{G} for $x \in \mathbb{R}^3$, generated by 5 different basis functions, thus $\alpha \in \Delta^5$. Optimal points of the individual basis functions, as well as the 1-level-sets of the quadratics are plotted, in order to give a sense of pairs $(Q_j, x_j^f)_{j=1}^5$ involved. The black edges connecting pairs of x_j^f correspond to edges of the geometric shape, but also to sets of global minima of pairwise function combinations. Compactness and path-connectedness are visually obvious for this example.

IV. GEOMETRY OF CONSTRAINED OPTIMAL SETS

This section will extend III to consider the case when $X \subseteq \mathbb{R}^n$ is a convex constraint set.

A. Convex Case

The following assumption and representation is required,

Assumption 3: There exists matrices $A_{eq} \in \mathbb{R}^{q \times n}$, $b_{eq} \in \mathbb{R}^q$ and C^1 convex functions $\{g_k(x)\}_{k=1}^r$ such that,

$$X = \{x \in \mathbb{R}^n \mid A_{eq}x = b_{eq}, g_k(x) \leq 0 \forall k = 1..r\}. \quad (14)$$

Assumption 4 (Slater's Condition): There exists a point $x' \in \mathbb{R}^n$ such that $A_{eq}x' = b$ and $g_k(x') < 0 \forall k = 1..n$, meaning X is non-empty.

Define $\mu \in \mathbb{R}_+^r$ and $\lambda \in \mathbb{R}^q$ as dual variables against the inequality and equality constraints describing X respectively. The Karush-Kuhn-Tucker (KKT) necessary conditions are sufficient to classify all optimal (minimizer) points of $f_\alpha(x)$ given a weighting $\alpha \in \Delta^n$ and assumptions 1, 3, 4 [22]:

$$\nabla_x f_\alpha(x) + A_{eq}^T \lambda + \sum_{k=1}^r \mu_k \nabla_x g_k(x) = 0 \quad (15a)$$

$$A_{eq}x = b \quad (15b)$$

$$g_k(x) \leq 0, \mu_k \geq 0 \quad \forall k = 1..n \quad (15c)$$

$$\sum_k \mu_k g_k(x) = 0. \quad (15d)$$

Define the symbol $\mathcal{W}^c = \mathbb{R}^n \times \Delta^n \times \mathbb{R}_+^r \times \mathbb{R}^q$ as the resident set containing $(x, \alpha, \mu, \lambda)$.

The optima-weight set in the constrained case is,

$$\hat{\mathcal{G}}^c = \{(x, \alpha, \mu, \lambda) \in \hat{\mathcal{W}} \mid \text{KKT conditions (15) hold}\} \quad (16)$$

$$\mathcal{G}^c = \pi^x \hat{\mathcal{G}}^c. \quad (17)$$

The feasibility LP to check if a $y \in \mathbb{R}^n$ is constrained-optimal (similar to (6) for the unconstrained case) is,

Problem 5 (Constrained Feasibility):

$$\text{find}_{\alpha \in \Delta^m, \mu \in \mathbb{R}_+^r, \lambda \in \mathbb{R}^q} (y, \alpha, \mu, \lambda) \in \hat{\mathcal{G}}^c, \quad (18)$$

with a constrained ProjGM program,

Problem 6 (Constrained ProjGM):

$$p^* = \min_{(x, \alpha, \mu, \lambda) \in \hat{\mathcal{G}}^c} \|y - x\|_2^2. \quad (19)$$

B. Quadratic Programming

We note that the special case of QP involves candidate functions and a constraint set,

$$f_j = x^T Q_j x / 2 + \varphi_j^T x, \quad \forall j = 1..m \quad (20a)$$

$$X = \{x \in \mathbb{R}^n \mid A_{eq}x = b_{eq}, Ax \leq b\}, \quad (20b)$$

for matrices $\{Q_j \in \mathbb{S}_+^n, \varphi_j \in \mathbb{R}^n\}_{j=1..m}$ and $A_{eq} \in \mathbb{R}^{q \times n}$, $b_{eq} \in \mathbb{R}^q$, $A \in \mathbb{R}^{r \times n}$, $b \in \mathbb{R}^r$ such that Assumption 4 (Slater) holds.

The constrained-optimal solution map given α and the parameters in (20) is,

$$\kappa^c(\alpha) = \text{argmin}_{x \in X} \sum_{j=1}^m \alpha_j f_j(x). \quad (21)$$

Remark 5: Due to possible weak convexity of some cost functions in \mathcal{F} , there may exist points $\alpha \in \Delta^m$ such that $\kappa^c(\alpha)$ is set-valued rather than single-valued. This will occur when $Q_\alpha = (\sum_{j=1}^m \alpha_j Q_j)$ is rank-deficient and $(\sum_{j=1}^m \alpha_j \varphi_j)$ is orthogonal to Q_α 's nullspace. In this case, the possibly discontinuous selection (translation of the minimum map),

$$s(\kappa^c(\alpha)) = \text{argmin}_{x \in \kappa^c(\alpha)} \|y - x\|_2^2, \quad (22)$$

will denote a constrained-optimal point in $\kappa^c(\alpha)$ that is closest to y . Finding a selection $s(\kappa^c(\alpha))$ requires solving a second QP over the X -intersected subspace of solutions of (21). This two-step approach involving a minimal selection was also performed in [23].

Figure 2 depicts constrained global optima \mathcal{G}^c , and it's unconstrained version \mathcal{G} generated from the same basis function set \mathcal{F} . Here 2 variables, 3 cost functions, and 4 inequality constraints with $(x, \alpha, \mu) \in \mathbb{R}^2 \times \Delta^3 \times \mathbb{R}_+^4$ are considered. Unconstrained and constrained optimal points of the individual basis functions, $x_{1:5}^f$ and $x_{1:5}^{f,c}$, are shown as colored dots. The black square denotes the feasible region. The blue ellipses are the level sets of the quadratic cost functions, chosen to highlight the spots where they are tangential to the constrained set boundary (implying the position of their constrained minima). The red box highlights the spot where the intersection of \mathcal{G}^c and the complement of \mathcal{G} is non-empty.

V. NUMERICAL METHODS

This section will present numerical approaches to find upper and lower bounds for the Unconstrained ProjGM 4 and Constrained ProjGM 6.

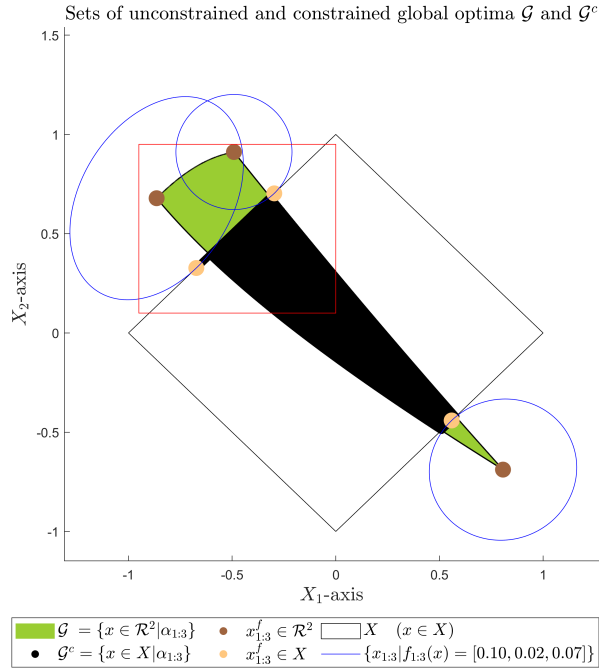


Fig. 2: The set \mathcal{G}^c is not necessarily a subset of \mathcal{G}

A. Upper Bounds

Upper bounds of p^* may be computed through sampling and local search optimization. The ProjGM problems may be formulated solely in terms of α through the use of the optimization maps $\kappa(\alpha)$ and $\kappa^c(\alpha)$ for $\alpha \in \Delta^*$ (bi-level optimization).

The objective in 4 and its gradient with respect to α is,

$$F(\alpha) = \|\kappa(\alpha) - y\|_2^2 \quad (23)$$

$$\nabla_{\alpha} F(\alpha)_j = 2(\kappa(\alpha) - y)^T \left(\sum_j \alpha_j Q_j \right)^{-1} Q_j (x_j^f - \kappa(\alpha)).$$

A Hessian for F may be similarly derived in closed form (omitted due to space constraints). Local solvers such as `fmincon` in MATLAB or `Manopt`¹ [24] may be given this value and derivative information to perform optimization over $\alpha \in \Delta^m$ (e.g. L-BFGS, trust-region).

The constrained objective in 6 may be expressed as,

$$F^c(\alpha) = \|s(\kappa^c(\alpha)) - y\|_2^2. \quad (24)$$

The constrained objective (24) is not generally differentiable with only the weak convexity assumption. Two possible options to minimize $F^c(\alpha)$ include gridding Δ^m , and using `fmincon` in terms of $(x, \alpha, \mu, \lambda) \in \hat{\mathcal{G}}^c$ directly on the constrained optimization problem 6.

B. Lower Bounds

Lower bounds $p_{low} \leq p^*$ to problems 4 and 6 in the quadratic case may be acquired through semidefinite programming. The presented method in this subsection is a

¹As Δ^m is not a manifold, the Hadamard parameterization $\Delta^m = \{z \circ z \mid \|z\|_2^2 = 1\}$ may be employed.

Quadratically Constrained Quadratic Programming (QCQP) that is equivalent to the degree-1 moment-Sum of Squares (SOS) hierarchy LMI enriched by additional constraints.

1) *Unconstrained Lower Bound*: Every point $(x, \alpha) \in \mathcal{W}$ defines a rank-1 PSD matrix $M = [1 \ x \ \alpha][1 \ x \ \alpha]^T$. Setting $M \in \mathbb{S}_+^{1+n+m}$ indexed by $(1, x, \alpha)$ as a matrix variable,

$$M = \begin{bmatrix} M_{11} & M_{1x} & M_{1\alpha} \\ M_{x1} & M_{xx} & M_{x\alpha} \\ M_{\alpha 1} & M_{\alpha x} & M_{\alpha\alpha} \end{bmatrix}, \quad (25)$$

the objective $\|y - x\|_2^2$ may be converted into an affine expression,

$$\sum_{i=1}^n (M_{x_i x_i} - 2y_i M_{1x_i}) + \|y\|_2^2 \quad (26)$$

Containment in $\hat{\mathcal{G}}$ from (9) may be expressed as,

$$\sum_{j=1}^m M_{\alpha_j 1} = 1 \quad (27a)$$

$$M_{\alpha_j 1} \geq 0 \quad \forall j = 1..m \quad (27b)$$

$$\sum_{j=1}^m Q_j M_{x\alpha_j} - (Q_j x_j^f) M_{1\alpha_j} = 0. \quad (27c)$$

Valid constraints are a set of relations that are always satisfied by an M generated by the optimal point (x^*, α^*) solving Problem 4. These valid constraints include,

$$M_{\alpha_i 1} = \sum_{j=1}^m M_{\alpha_i \alpha_j} \quad \forall i \neq j \quad (28a)$$

$$M_{\alpha_i \alpha_j} \geq 0 \quad \forall i \neq j \quad (28b)$$

$$M_{\alpha_i \alpha_j} \leq \alpha_i, \quad M_{\alpha_i \alpha_j} \leq \alpha_j \quad \forall i \neq j \quad (28c)$$

$$M_{\alpha_i \alpha_i} \leq M_{1\alpha_i} \quad \forall i = 1..m \quad (28d)$$

$$M_{\alpha_i \alpha_j} \leq 1/4 \quad \forall i \neq j. \quad (28e)$$

Constraints (28a)-(28b) arise from multiplying together defining constraints for the simplex. The diagonal entry $M_{\alpha_i \alpha_i}$ will automatically be positive by $M \in \mathbb{S}_+^{1+n+m}$. Constraints (28c)-(28e) originate from observations about the simplex Δ^m . Because $\alpha \in \Delta^m \subseteq [0, 1]^m$, coordinate-wise multiplication will satisfy $\alpha_i \alpha_j \leq \alpha_i$ for all $i, j \in 1..n$.

Constraint (28e) results from the fact that elementary symmetric polynomials on the probability simplex Δ^σ are maximized at the vector $\mathbf{1}/\sigma$. The elementary symmetric polynomial applied to $\alpha' \in \Delta^\sigma$ is $e_2^\sigma(\alpha') = \sum_{1 \leq i < j \leq n} \alpha'_i \alpha'_j$, and admits the maximum value of $\max e_2^\sigma = \binom{\sigma}{2} \frac{1}{\sigma^2}$. The valid inequality derived from $\sigma = 2$ with $\max e_2^\sigma = \frac{1}{4}$ is written in (28e). Valid inequalities with higher σ may be written at the cost of including a combinatorially increasing number of constraints.

Algorithm 1: Unconstrained LMI

Input: y, Q, x_f

Output: p_{low}^* , M (or Infeasibility)

Solve (or find infeasibility certificate):

$$p_{low}^* = \min_M \text{Objective (26)} \quad (29a)$$

$$\text{Optimality (27), Valid (28)} \quad (29b)$$

$$M_{11} = 1, \quad M \in \mathbb{S}_+^{1+n+m}. \quad (29c)$$

A rank-1 matrix solution M of Algorithm 1 certifies that $p_{low}^* = p^*$. The optimal entries (x, α) can then be read from the solution's entries $(M_{1x}, M_{1\alpha})$. Adding valid inequalities (28) can encourage rank-1 solutions of LMI lower bound problem, refer to [25] for further examples of this phenomenon.

2) *Constrained Lower Bounds*: The lower bound SDP for the QP setting in Section IV-B requires a matrix $M \in \mathbb{S}_+^{1+n+m+r}$ and a vector $\lambda \in \mathbb{R}^q$ such that the entries of M are indexed by $[1, x, \alpha, \mu]$. The equality multipliers λ may be omitted from M because there is no multiplication in (15) between terms that contain x and λ . The affine constraint interpretation of the KKT conditions in (15) is,

$$\sum_{j=1}^m (Q_j M_{x\alpha_j} + \varphi_j M_{\alpha_j 1}) + A_{eq}^T \lambda + A^T M_{\mu 1} = 0 \quad (30a)$$

$$A_{eq} M_{x1} = b_{eq} \quad (30b)$$

$$A M_{x1} \leq b \quad M_{\mu 1} \geq 0 \quad (30c)$$

$$-b^T M_{\mu 1} + \sum_{k=1}^r A_k M_{x\mu_k} = 0. \quad (30d)$$

The notation A_k in the complementary slackness constraint (30d) indicates row k of the matrix A . Valid inequalities for the constrained QP case include (with inequality constraint indices (k, ℓ)),

$$M_{\mu_k \mu_\ell} \geq 0 \quad \forall k \neq \ell \quad (31a)$$

$$M_{\mu_k \alpha_i} \geq 0 \quad \forall k = 1..r, i = 1..m. \quad (31b)$$

$$M_{\mu_k \alpha_i} \leq M_{\mu_k 1} \quad \forall k = 1..r, i = 1..m. \quad (31c)$$

Both $\alpha \in \Delta^m$ and $\mu \in \mathbb{R}_+^r$ are nonnegative, so their multiplications should also be nonnegative. The QP-lower-bounding LMI is,

Algorithm 2: Constrained LMI

Input: $y, Q, \varphi, A, b, A_{eq}, b_{eq}$

Output: p_{low}^*, M (or Infeasibility)

Solve (or find infeasibility certificate):

$$p_{low}^* = \min_{\lambda \in \mathbb{R}^q, M} \text{Objective (26)} \quad (32a)$$

$$\text{Simplex (27a)-(27b), KKT (30)} \quad (32b)$$

$$\text{Valid (28), (31)} \quad (32c)$$

$$M_{11} = 1, M \in \mathbb{S}_+^{1+n+m+r} \quad (32d)$$

The accuracy of $p_{low}^* \leq p^*$ from Algorithm 2 may be improved if an upper bound for μ_k^2 ($M_{\mu_k \mu_k}$) was known for each k . Bisection-based approaches with convex cost heuristics may also be applied to the presented LMIs [26].

VI. NUMERICAL EXAMPLE

This section contains a small-scale toy example with visuals, to develop intuition. The example will showcase flaws in the approximate projection method [9] when an inconsistent basis is supposed, and the robustness of the bi-level method. The proposed distance-bounding methods will also be displayed. Future works will showcase the method for larger-scale examples and will be applied to real data.

The method is designed to invalidate a proposed set of basis functions with respect to a data set. In other words, the method is able to say whether IOC with proposed basis functions can be successfully applied to a given data set.

Assume that an observed human decision (or trajectory) can be represented by a vector $y \in \mathbb{R}^2$, and that it is generated via an unknown process (as is actually the case in real-world applications). Assume a basis \mathcal{F} of 5 quadratic cost functions believed to be underlying the decision-making (or trajectory-generation). Figure 3 depicts the 2-dimensional test point $y \in \mathbb{R}^2$, and the set of unconstrained global optima \mathcal{G} which is generated using \mathcal{F} .

On Figure 3 the actual minimum distance projection $\text{Proj}_{\mathcal{G}}(y)$ is shown as reference, and is computed by a combined brute-force grid search and local bi-level search on the cost function parametrization $\alpha \in \Delta^5$. The minimum distance local bi-level formulation result, initialized with $\alpha_i = 0.2$, is annotated as ProjGM and shown, alongside the Keshavaraz minimum KKT-constraint violation formulation [9]. The M_{x1} entry of the PSD matrix output by algorithm 1 is also shown.

Projection of a random point onto unconstrained set of global optima \mathcal{G}

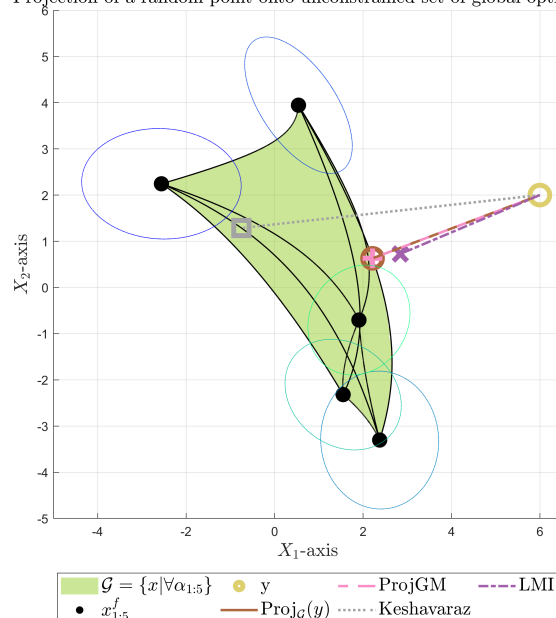


Fig. 3: Comparison of the projection onto an unconstrained set of optima \mathcal{G} .

From figure 3 it is possible to see that the approximate Keshavaraz method can impute a strongly inconsistent objective if the test point is outside the set of global optima \mathcal{G} . In this example the local bi-level search reaches the global minimum of distance, but may get stuck in a local minimum depending on problem structure.

Given a data set of a 100 points $y \in \mathbb{R}^2$ generated by adding Gaussian noise to the shown test point in 3 we calculate the distance of each of them to the set of global optima with different methods. We obtain that the mean distance to the set of global optima for different methods

is, in ascending order, $d_{\text{LMI}} = 3.78$, $d_{\text{ProjGM}} = 4.09$ and $d_{\text{Keshavaraz}} = 6.98$, the true value being $d_{\text{ProjGM}} = 4.09$.

In practice, for higher-dimensional data, running the bilevel method may be expensive. If the Keshavaraz method does not produce a zero or extremely-close-to-zero error, it is useful to be able to run the LMI algorithm to check the lower-bound of the distance from the set of global optima. If the lower bound is higher than the desired error in data replication, we can conclude that the proposed set of basis functions \mathcal{F} is not good enough to represent the data.

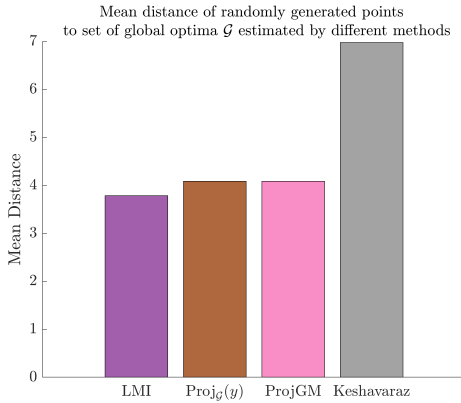


Fig. 4: Comparison of the projection distance onto an unconstrained set of optima \mathcal{G} .

MATLAB (2021b) code to replicate figures and experiments is publicly available at <https://github.com/jarmill/inverse-optimal>. Dependencies include Mosek [27] and YALMIP [28].

VII. EXTENSIONS

This section will detail extensions of the current work and some discussion of future research directions.

A. Piecewise Functions and Convex Lifts

The method of convex lifts [29, 30] may be applied to solve the ProjGM problem over some classes of piecewise-defined costs f_j . Assume there exists a set of C^1 -continuous functions $\{w_{j\ell}(x)\}_{\ell=1}^{L_j}$ with L_j finite such that the convex f_j satisfies,

$$f_j(x) = \max_{\ell \in 1..L_j} w_{j\ell}(x). \quad (33)$$

Remark 6: Convexity is preserved under the pointwise maximum operation (though not under pointwise minimum).

Remark 7: The class of convex piecewise affine functions with $w_{j\ell}(x) = m_{j\ell}^T x - b_{j\ell}$ for $m_{j\ell} \in \mathbb{R}^{n \times 1}$, $b_{j\ell} \in \mathbb{R}$ may be expressed as an instance of (33).

New variables τ_j may be added for every max-representable function $f_j(x)$, forming the equivalent optimization problems with equal objectives:

$$f^* = \min_{x \in X} \sum_{j=1}^m f_j(x) \quad (34)$$

$$f^* = \min_{x \in X, \tau \in \mathbb{R}^m} \sum_{j=1}^m \tau_j \quad (35)$$

$$\alpha_j w_{j\ell}(x) \leq \tau_j \quad \forall \ell = 1..L_j, j = 1..m.$$

The formerly non-differentiable objectives $f_j(x)$ from (34) are lifted into the constrained problem (35) where each $w_{j\ell}(x)$ is differentiable for all $x \in X$. KKT equations may then be written for (35) and then utilized in (16) to describe the cone $\hat{\mathcal{G}}^c(x, \tau, \alpha, \mu, \lambda)$ for use in the constrained ProjGM problem 6.

Remark 8: Minimization of the L_1 norm $\|x\|_1$ may be expressed as $\min \sum_{i=1}^n t_i : -t_i \leq x_i \leq t_i$ in $2n$ inequality constraints by adding n new variables $\{t_i\}_{i=1}^n$ for use in constrained ProjGM. L_p norms with rational $p \in [1, \infty)$ admit second-order-cone representations [31] through lifting, and can therefore be members of the dictionary \mathcal{F} .

A rank-1 solution M of (29) certifies that $p_{\text{low}}^* = p^*$ from (7), and the optimal (x, α) may be read from M_{x1} and $M_{\alpha 1}$.

B. Projection onto Local Minimizers

This paper was restricted to convex cost functions in \mathcal{F} and convex sets X . Problem 2 may be extended to finding the minimum distance p^* between y and some local minimizer of $f_\alpha(x)$. Theorem 12.6 (Eq. (12.65)) of [32] outlines second-order necessary conditions for a point x^* to be a local minimizer. In the unconstrained case, the Hessian matrix $\nabla_x^2 f_\alpha(x)$ must be PD in addition to (x, α) satisfying the first-order condition $\nabla_x f_\alpha(x) = 0$. The constrained case with non-convex functions requires more delicacy, as the quadratic form $w \rightarrow w^T \nabla_x^2 f_\alpha(x) w$ must be positive for all $w \neq 0$ vectors inside the Critical Cone (Eq. (12.53) of [32]) formed by (x, λ, μ) .

C. Polynomial Optimization

The ProjGM programs 4 and 6 are instances of Polynomial Optimization Problems when $f_j(x), g_k(x)$ are all convex polynomial functions of x . The sets $\hat{\mathcal{G}}(x, \alpha), \hat{\mathcal{G}}^c(x, \alpha, \mu, \lambda)$ are basic semialgebraic sets, and their projections $\mathcal{G}(x), \mathcal{G}^c(x)$ are in turn semialgebraic sets. The projections $\mathcal{G}(x), \mathcal{G}^c(x)$ may be analyzed by quantifier elimination algorithms such as the Cylindrical Algebraic Decomposition [33]. The moment-SOS hierarchy is a method that yields a rising sequence of lower bounds to the distances p^* by solving a sequence of LMIs in combinatorially increasing size [34]. Theorem 3.2 assures (under mild conditions) that the moment-SOS relaxations of the unconstrained problem 4 will converge at a finite degree. Convergence of the moment-SOS hierarchy for bilevel Polynomial Optimization Problems are established in [35]. The LMI presented in Section V-B is an instance of the degree-1 moment-SOS hierarchy as enriched with valid constraints.

Given arbitrary convex polynomials $f_j(x), g_k(x)$, all constraints in $\hat{\mathcal{G}}(x, \alpha)$ are affine w.r.t. α . Likewise, the describing constraints of $\hat{\mathcal{G}}^c(x, \alpha, \mu, \lambda)$ are affine in (α, μ, λ) all together. A theorem of alternatives may be used to eliminate the affine-dependent groups (α) or (α, μ, λ) , yielding a set of linked LMI constraints of smaller size that solely depend on x [36]. This reduction in the number of variables decreases the computational burden of solving LMIs as the degree increases. A full presentation and application of polynomial optimization for ProjGM will take place in sequel work.

VIII. CONCLUSION

This paper introduced a cost functions set invalidation interpretation of the ProjGM problem. The geometry of optimal sets were explored in the constrained and unconstrained cases (focusing on quadratics). Numerical algorithms were implemented to upper and lower bound the ProjGM optima. Future work includes performing cost-function discovery in order to generate candidate functions $f(x)$ that would reduce the distance p^* from (4) when added to the set of cost functions \mathcal{F} . The proposed method should be applied for IOC problem based actual noisy human data.

On the theoretical side, sensitivity methods from set valued analysis will be applied to determine which properties of \mathcal{G} and $\hat{\mathcal{G}}$ are preserved when assumptions are lifted (e.g. does Theorem 3.3 hold when all f_j are strongly convex rather than strongly convex quadratic?). Continuity properties of the solution map (21) and its selection (22) will also be explored [37, 38].

ACKNOWLEDGEMENTS

F. Bečanović would like to thank Florent Lamiroux for his useful remarks. J. Miller would like to thank his advisors Mario Sznaiar and Didier Henrion, and the POP group at LAAS-CNRS for their comments and support.

REFERENCES

- [1] B. Berret, E. Chiovetto, F. Nori, and T. Pozzo, "Evidence for composite cost functions in arm movement planning: an inverse optimal control approach," *PLoS computational biology*, vol. 7, no. 10, p. e1002183, 2011.
- [2] S. Albrecht, M. Leibold, and M. Ulbrich, "A bilevel optimization approach to obtain optimal cost functions for human arm movements," *Numerical Algebra, Control & Optimization*, vol. 2, no. 1, p. 105, 2012.
- [3] 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, 2010.
- [4] J. F.-S. Lin, V. Bonnet, A. M. Panchea, N. Ramdani, G. Venture, and D. Kulić, "Human motion segmentation using cost weights recovered from inverse optimal control," in *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2016, pp. 1107–1113.
- [5] A. M. Panchea, N. Ramdani, V. Bonnet, and P. Fraisse, "Human Arm Motion Analysis Based on the Inverse Optimization Approach," in *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*. IEEE, 2018, pp. 1005–1010.
- [6] K. Westermann, J. F.-S. Lin, and D. Kulić, "Inverse optimal control with time-varying objectives: application to human jumping movement analysis," *Scientific reports*, vol. 10, no. 1, pp. 1–15, 2020.
- [7] D. Clever and K. D. Mombaur, "An Inverse Optimal Control Approach for the Transfer of Human Walking Motions in Constrained Environment to Humanoid Robots," in *Robotics: Science and systems*, 2016.
- [8] A. Aswani, Z.-J. Shen, and A. Siddiq, "Inverse Optimization with Noisy Data," *Operations Research*, vol. 66, no. 3, pp. 870–892, 2018.
- [9] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a Convex Objective Function," in *2011 IEEE international symposium on intelligent control*. IEEE, 2011, pp. 613–619.
- [10] W. Jin, D. Kulić, J. F.-S. Lin, S. Mou, and S. Hirche, "Inverse optimal control for multiphase cost functions," *IEEE Transactions on Robotics*, vol. 35, no. 6, pp. 1387–1398, 2019.
- [11] P. Englert, N. A. Vien, and M. Toussaint, "Inverse kkt: Learning cost functions of manipulation tasks from demonstrations," *The International Journal of Robotics Research*, vol. 36, no. 13-14, pp. 1474–1488, 2017.
- [12] W. Jin, Z. Liang, and S. Mou, "Inverse optimal control from demonstration segments," *arXiv preprint arXiv:2010.15034*, 2020.
- [13] W. Jin, D. Kulić, S. Mou, and S. Hirche, "Inverse optimal control from incomplete trajectory observations," *The International Journal of Robotics Research*, vol. 40, no. 6-7, pp. 848–865, 2021.
- [14] A. M. Panchea and N. Ramdani, "Towards solving inverse optimal control in a bounded-error framework," in *2015 American Control Conference (ACC)*. IEEE, 2015, pp. 4910–4915.
- [15] —, "Inverse parametric optimization in a set-membership error-in-variables framework," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6536–6543, 2017.
- [16] J. Colomel, D. Daney, and F. Charpillat, "On the Reliability of Inverse Optimal Control," 2021.
- [17] M. Johnson, N. Aghasadeghi, and T. Bretl, "Inverse optimal control for deterministic continuous-time nonlinear systems," in *52nd IEEE Conference on Decision and Control*. IEEE, 2013, pp. 2906–2913.
- [18] J. Thai and A. M. Bayen, "Imputing a Variational Inequality Function or a Convex Objective Function: a Robust Approach," *Journal of Mathematical Analysis and Applications*, vol. 457, no. 2, pp. 1675–1695, 2018.
- [19] R. Stagni, A. Leardini, A. Cappozzo, M. G. Benedetti, and A. Cappello, "Effects of hip joint centre mislocation on gait analysis results," *Journal of biomechanics*, vol. 33, no. 11, pp. 1479–1487, 2000.
- [20] S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnykova, "Bilevel Programming Problems," *Energy Systems*. Springer, Berlin, vol. 10, pp. 978–3, 2015.
- [21] J. R. Munkres, "Topology," 2000.
- [22] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.
- [23] J. Spjøtvold, P. Tøndel, and T. Johansen, "Continuous Selection and Unique Polyhedral Representation of Solutions to Convex Parametric Quadratic Programs," *Journal of Optimization Theory and Applications*, vol. 134, no. 2, pp. 177–189, 2007.
- [24] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *Journal of Machine Learning Research*, vol. 15, no. 42, pp. 1455–1459, 2014. [Online]. Available: <https://www.manopt.org>
- [25] A. A. Ahmadi and J. Zhang, "Semidefinite Programming and Nash Equilibria in Bimatrix Games," *INFORMS Journal on Computing*, vol. 33, no. 2, pp. 607–628, 2021.
- [26] K. Mohan and M. Fazel, "Iterative Reweighted Algorithms for Matrix Rank Minimization," *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 3441–3473, 2012.
- [27] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.2.*, 2020. [Online]. Available: <https://docs.mosek.com/9.2/toolbox/index.html>
- [28] J. Lofberg, "YALMIP : a toolbox for modeling and optimization in MATLAB," in *ICRA (IEEE Cat. No.04CH37508)*, 2004, pp. 284–289.
- [29] M. Yannakakis, "Expressing combinatorial optimization problems by Linear Programs," *Journal of Computer and System Sciences*, vol. 43, no. 3, pp. 441–466, 1991.
- [30] J. Gouveia, P. A. Parrilo, and R. R. Thomas, "Lifts of Convex Sets and Cone Factorizations," *Mathematics of Operations Research*, vol. 38, no. 2, pp. 248–264, 2013.
- [31] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Mathematical programming*, vol. 95, no. 1, pp. 3–51, 2003.
- [32] J. Nocedal and S. Wright, "Numerical optimization," *Springer Science*, vol. 35, pp. 67–68, 1999.
- [33] B. F. Caviness and J. R. Johnson, *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Springer Science & Business Media, 2012.
- [34] J. B. Lasserre, *Moments, Positive Polynomials And Their Applications*, ser. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009.
- [35] V. Jeyakumar, J. B. Lasserre, G. Li, and T. Pham, "Convergent Semidefinite Programming Relaxations for Global Bilevel Polynomial Optimization Problems," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 753–780, 2016.
- [36] A. Ben-Tal, D. Den Hertog, and J.-P. Vial, "Deriving robust counterparts of nonlinear uncertain inequalities," *Mathematical programming*, vol. 149, no. 1-2, pp. 265–299, 2015.
- [37] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*. Springer Science & Business Media, 2009.
- [38] N. N. Tam, "On continuity properties of the solution map in quadratic programming," *Acta Mathematica Vietnamica*, vol. 24, pp. 47–61, 1999.