

# Convex Formulations and Algebraic Solutions for Linear Quadratic Inverse Optimal Control Problems

Marcel Menner and Melanie N. Zeilinger

**Abstract**—This paper presents convex formulations for inverse optimal control problems for linear systems to infer cost function matrices of a quadratic cost from both optimal and non-optimal closed-loop gains. It introduces an optimality measure which enables a formulation of the problem as a convex semidefinite program for the general case and a linear program for several special cases. We derive an explicit algebraic expression for general objective function matrices as well as conditions under which the solution to the inverse optimal control problem is unique. The result is derived by means of a vectorization and parametrization of the algebraic Riccati equation. A simulation example highlights the robust performance in the presence of noise on the measured closed-loop gain and the computational efficiency of the proposed problem formulations.

## I. INTRODUCTION

Inverse optimal control (IOC) addresses the problem of inferring the cost function matrices of a corresponding unconstrained optimal control problem from a given control law. It provides a promising approach for learning from observed behavior, offering favorable generalization properties by relating actions to an underlying objective function. While observed data generally only provides sparse information about the control law, the objective function allows for generating a control law for the entire state space. One application motivating this paper is personalized learning with the goal of tailoring system operation to match user-specific demands. The challenges in this context are the generalization of a user’s intention from observed behavior and online processing of the collected data. The IOC techniques proposed in this paper address both challenges by introducing efficient solutions for recovering cost functions from both optimal and non-optimal closed-loop gains based on convex optimization problems. While the presented methods can be applied in the context of imitation learning or identifying preferences in user-operated systems, the techniques address general IOC problems.

Optimal control is well-studied in the literature and, for linear systems and quadratic cost functions, there exists an algebraic solution to the unconstrained infinite-horizon control problem via the algebraic Riccati equation (ARE), e.g. [1]. IOC distinguishes itself from other learning solutions such as reinforcement learning, e.g. [2], inverse reinforcement learning, e.g. [3]–[5], or apprenticeship learning, e.g. [6]–[8], by means of its algebraic connection between the cost function and the control law, i.e. the ARE.

M. Menner and M. N. Zeilinger are with the Institute for Dynamic Systems and Control, ETH Zürich, 8092 Zürich, Switzerland {mmenner, mzeilinger}@ethz.ch

This work was supported by the Swiss National Science Foundation under grant no. PP00P2 157601 / 1.

This algebraic connection is exploited in the IOC techniques in [9], [10] and inverse model predictive control (MPC) in [11]. In [9]–[11], a semidefinite program (SDP) is proposed to infer the objective function from an optimal closed-loop gain and a bilinear SDP is solved with a gradient descent method for a non-optimal closed-loop gain. Non-optimal closed-loop gains, however, are expected to be commonly encountered in practice, since the gain will be recovered from noisy input data. Other related work can be found in [12]–[14]. In [12], it is shown that, for a linear system, any feedback gain is optimal for some choice of cost matrices if the cost structure includes terms with state-input coupling. Related work in economic MPC [13] considers the problem of computing positive definite cost matrices, which yield the same feedback gain as an indefinite economic cost. The method proposed in [13] also solves an IOC problem with the techniques in [12], however, the objective is to keep matrices well-conditioned, rather than reproduce optimal behavior.

This paper makes the following contributions: It introduces convex formulations for the IOC problem for linear systems and a quadratic cost by means of an optimality measure reflecting the infinite-horizon cost, which has the benefit of offering a notion of closed-loop performance. More specifically, we present an SDP formulation for inferring general objective function matrices and a linear programming (LP) formulation for both diagonal objective function matrices and block-diagonal matrices with blocks of dimension two. Different from [9]–[11], this paper shows that the IOC problem can be stated as a convex optimization problem for both optimal and non-optimal closed-loop gains. The results offer an alternative approach to [12] by making use of a standard quadratic cost function without the requirement for introducing state-input coupling. In [12], noisy measurements necessarily lead to a nonzero state-input coupling term in the cost, whereas in the proposed approach an objective function with only quadratic terms on states and inputs is inferred which best explains the observed measurements. Furthermore, an explicit algebraic expression for the inverse problem given an optimal controller gain is derived and sufficient conditions are stated under which the corresponding cost function can be uniquely inferred for diagonal cost matrices. Finally, we show feasibility of the convex formulations for general, i.e. in particular non-optimal gains, for all cases. The performance of the proposed solutions is analyzed in simulations, where the cost function is inferred from noisy measurements of the feedback gain with various standard deviations.

The paper is structured as follows. In Section II, the

problem is stated. Section III presents the notation and preliminaries. Section IV introduces the optimality measure and derives the convex optimization problems. Section V derives the explicit algebraic solution to the inverse problem and states its properties. A simulation is presented in Section VI and the paper is concluded with Section VII.

## II. PROBLEM STATEMENT

Consider a linear discrete-time system of the form

$$x(k+1) = Ax(k) + Bu(k),$$

where  $x(k)$  and  $u(k)$  are the state and input at time instance  $k$ , respectively, and  $A \in \mathbb{R}^{n \times n}$  and  $B \in \mathbb{R}^{n \times m}$  are the state and input matrix, respectively.

We address the inverse optimal control problem, i.e. inferring the generating optimal control problem from an observed control law, for linear systems and quadratic cost functions. Control actions are modeled as the solution of an unconstrained infinite-horizon optimal control problem of the following form:

$$\begin{aligned} x(k)^\top Px(k) = \min \sum_{i=0}^{\infty} (x_i^\top Qx_i + u_i^\top Ru_i) \\ \text{s.t. } x_{i+1} = Ax_i + Bu_i \\ x_0 = x(k), \end{aligned} \quad (1)$$

where  $P \succ 0$  describes the infinite-horizon cost,  $Q \succeq 0$  is the state-penalty matrix, and  $R \succ 0$  is the input-penalty matrix [1]. The task is to infer the cost function matrices from an observed control law  $u(k) = Kx(k)$ . In contrast to [9]–[12], we propose to optimize for  $Q$  and  $R$ , which minimize violation of the optimality conditions for the actual measured  $K$ . As we will show in the following sections, this enables a convex formulation to the IOC problem. Further, we will present an explicit algebraic solution to the inverse problem of (1). The results are based on the following assumptions:

**Assumption 1.** *The system matrix  $A$  and input matrix  $B$  are given and the state  $x(k)$  and input  $u(k)$  can be measured.*

**Assumption 2.**  *$(A, B)$  is stabilizable and  $(Q^{1/2}, A)$  is detectable.*

**Assumption 3.**  *$B$  has full column rank.*

**Assumption 4.** *All eigenvalues of the closed-loop system  $A + BK$  are nonzero, i.e.  $(A + BK)^{-1}$  exists.*

**Remark 1** (Existence of  $(A + BK)^{-1}$ ). *For a physical system, Assumption 4 is generally satisfied. If, however, one encounters singular  $A + BK$ , e.g. due to noise, a subsystem  $A_s = \Xi(A + BK)\Xi^T$  can be considered to infer cost function matrices  $Q_s$  and  $R$ . The projection  $\Xi \in \mathbb{R}^{n_p \times n}$  is defined such that  $\Xi(A + BK)\Xi^T$  has full rank and  $\Xi v_i = 0$ , where  $v_i$  with  $i, \dots, n - n_r$  are the eigenvectors of  $A + BK$ , which correspond to zero eigenvalues.  $n_r$  is the dimension of zero eigenvalues. The state-penalty matrix is then obtained as  $Q = \Xi^T Q_s \Xi$ .*

**Remark 2** (Derivation of controller gain). *We will show how to infer  $Q$  and  $R$  from a given controller gain  $K$ . The gain*

*can be obtained from input and output measurements with standard approaches, such as least squares techniques [15].*

## III. NOTATION & PRELIMINARIES

### A. Matrix notation

We denote with  $\mathbb{1}$  a column vector of ones of appropriate dimension and with  $I_n \in \mathbb{R}^{n \times n}$  an identity matrix. Let  $\mathbb{I}$  be an index set. Then  $q_{\mathbb{I}}$  selects the entries of a vector  $q$  indexed by  $\mathbb{I}$  and  $M_{*\mathbb{I}}$  selects the columns of a matrix  $M$  indexed by  $\mathbb{I}$ . We denote with  $\text{diag}(M)$  a column vector consisting of the diagonal elements of a matrix  $M$  and with  $\text{diag}(m)$  a diagonal matrix consisting of the elements of a vector  $m$ .  $M_{ij}$  refers to the element of matrix  $M$  associated with row  $i$  and column  $j$ .

### B. Vectorized notation

Consider the matrix equation

$$XYZ = 0 \quad (2)$$

where  $X \in \mathbb{R}^{l \times m}$ ,  $Y \in \mathbb{R}^{m \times n}$ , and  $Z \in \mathbb{R}^{n \times p}$ . The vec-operator is defined as

$$y := \text{vec}(Y) = \begin{bmatrix} Y_{*1} \\ Y_{*2} \\ \vdots \\ Y_{*n} \end{bmatrix},$$

where  $y \in \mathbb{R}^{mn \times 1}$ . It follows that (2) is equivalent to

$$\text{vec}(XYZ) = 0 \Leftrightarrow (Z^\top \otimes X) \text{vec}(Y) = 0$$

with the Kronecker product  $\otimes$  [16]. In order to solve a system of equations

$$XYZ + CY^\top D = 0$$

for  $Y^\top$ , we can therefore reformulate the condition in vectorized form, i.e.

$$\begin{aligned} (Z^\top \otimes X) \text{vec}(Y) + (D^\top \otimes C) \text{vec}(Y^\top) &= 0 \\ \Leftrightarrow ((Z^\top \otimes X) + (D^\top \otimes C)T) \text{vec}(Y) &= 0 \end{aligned}$$

with the permutation  $T \in \mathbb{R}^{mn \times mn}$  such that

$$\text{vec}(Y^\top) = T \text{vec}(Y). \quad (3)$$

### C. Algebraic conditions for optimal control

The following well-known results from optimal control exploiting the algebraic relationship between cost and controller gain are used as a basis for the inverse optimal control solutions derived in this paper. The optimal feedback gain  $K$  is defined as

$$\begin{aligned} K &= -(B^\top PB + R)^{-1} B^\top PA \\ \Leftrightarrow B^\top P(A + BK) &= -RK, \end{aligned} \quad (4)$$

where  $P$  as in (1) is the infinite-horizon cost and the solution of the discrete-time algebraic Riccati equation, e.g. [1]:

$$\begin{aligned} P &= Q + A^\top PA - A^\top PB (B^\top PB + R)^{-1} B^\top PA \\ &= Q + A^\top P(A + BK). \end{aligned} \quad (5)$$

#### IV. CONVEX FORMULATIONS OF THE INVERSE OPTIMAL CONTROL PROBLEM

Based on the ARE in (5), the following introduces an SDP formulation for the IOC problem with general objective function matrices, two LP formulations for recovering block-diagonal and diagonal objective function matrices, and an explicit algebraic solution for recovering diagonal objective function matrices from general feedback gains  $K$ , which may be non-optimal. In general, gains are expected to be non-optimal with respect to the optimal control problem (1), since they are recovered from noisy input data. In this case, the optimality condition (5) is infeasible and cannot be satisfied for any  $Q$ ,  $R$ , and  $P$ .

##### A. Cost function & Parametrization

The key for providing a convex IOC formulation is the use of a particular objective, which will be introduced in the following. The objective measures non-optimality as violation of the optimality condition (5). Hence, we aim at finding  $Q$ ,  $R$ , and  $P$  that result in the smallest deviation  $\Lambda$  from the infinite-horizon cost  $P$  satisfying the optimality condition:

$$\begin{aligned} (P + \Lambda) &= Q + A^\top (P + \Lambda)(A + BK) \Leftrightarrow \\ \Lambda - A^\top \Lambda (A + BK) &= Q - P + A^\top P (A + BK). \end{aligned} \quad (6)$$

By the interpretation of  $P$  as the optimal infinite horizon cost, the deviation  $\Lambda$  also offers a notion of closed-loop performance, i.e. the gain  $K$  is optimal if and only if  $\Lambda = 0$ . This intuitive optimality measure is different from the literature, e.g. [9]–[11], generally minimizing the distance between the measured and an optimal feedback gain. Note that we neither enforce positive definiteness nor symmetry of  $P$  in the following IOC problems. However,  $P$  will be positive definite if  $\Lambda = 0$  and symmetry is encouraged by the objective.

We first introduce a parametrization that allows for the algebraic solution and uniqueness statements in Section V. Since there are no constraints on the structure of  $P$ , it is possible to substitute  $P$  in (6) with an expression derived from (4):

$$P = -B^+ RK (A + BK)^{-1} + B^{\perp\top} Y, \quad (7)$$

where  $B^+ := B(B^\top B)^{-1}$ ,  $Y \in \mathbb{R}^{(n-m) \times n}$ , and  $B^\perp \in \mathbb{R}^{(n-m) \times n}$  with  $B^\perp B = 0$  is a basis for the orthogonal complement of  $B$ . Assumption 3 ensures the existence of  $(B^\top B)^{-1}$ . Condition (6) can then be replaced by

$$\begin{aligned} \Lambda - A^\top \Lambda (A + BK) &= Q + B^+ RK (A + BK)^{-1} \\ &\quad - B^{\perp\top} Y - A^\top B^+ RK \\ &\quad + A^\top B^{\perp\top} Y (A + BK). \end{aligned} \quad (8)$$

For consistency of notation in the paper we present all results using the parametrization. For general cost function matrices, however, one could directly use (6) and optimize over  $P$ . Instead of solving (4) and (6) for  $Q$ ,  $R$ , and  $P$ , the parametrization allows to solve (8) for  $Q$ ,  $R$ , and  $Y$ . While the formulations are equivalent, the parametrization in

$Y$  has the advantage of reducing the number of optimization variables and constraints both by  $nm$ .

In addition to measuring violation of optimality, the objective penalizes non-symmetry of the matrix  $P$  by means of the measure  $\Lambda_p = P - P^\top$ , i.e.

$$\begin{aligned} \Lambda_p &= -B^+ RK (A + BK)^{-1} + B^{\perp\top} Y \\ &\quad + \left( K (A + BK)^{-1} \right)^\top R (B^+)^{\top} - Y^\top B^\perp. \end{aligned} \quad (9)$$

Notice that  $\Lambda_p = 0$ , i.e.  $P$  is symmetric, if the measured  $K$  is the optimal solution of the problem in (1). Eq. (9) introduces a preference on  $R$  and  $Y$  that render  $P$  symmetric.

**Remark 3.** Without loss of generality, we can choose  $R_{11} = 1$  as only the relative weighting between  $Q$  and  $R$  is relevant for the optimal solution [1]. A diagonal element of  $R$  is used to fix the scaling of the objective function as  $R \succ 0$ , whereas  $Q \succeq 0$  and diagonal elements of  $Q$  may be zero.

**Remark 4.** The formulations presented in the following can similarly be derived for continuous-time systems of the form  $\dot{x}(t) = A_c x(t) + B_c u(t)$ .

##### B. SDP formulation for general cost matrices

First, an SDP formulation for inferring general objective function matrices from general gains  $K$  is proposed. Given  $K$ , the goal is to find  $Q$ ,  $R$ , and  $Y$ , which minimize the deviation to optimality  $\Lambda$  in (8) and to symmetry  $\Lambda_p$  in (9).

The SDP for a general linear quadratic IOC problem is

$$\begin{aligned} \min \|\Lambda\|_1 + c_p \cdot \|\Lambda_p\|_1 \\ \text{s.t. } & (8) \\ & (9) \\ & Q \succeq 0 \\ & R \succeq \epsilon \cdot I_m \\ & R_{11} = 1 \end{aligned} \quad (10)$$

with  $c_p \geq 0$  as remaining design parameter for trading-off optimality against symmetry and  $\epsilon > 0$  to ensure  $R \succ 0$ .

##### C. LP formulation for diagonal cost matrices

The most common objective function in optimal control problems makes use of diagonal cost matrices, e.g. [1]. This motivates the derivation of an LP formulation for reconstructing diagonal cost matrices, which can be solved efficiently with available solvers such as CPLEX [17].

The LP for inferring diagonal cost matrices is given by

$$\begin{aligned} \min \|\Lambda\|_1 + c_p \cdot \|\Lambda_p\|_1 \\ \text{s.t. } & (8) \\ & (9) \\ & \text{diag}(Q) \geq 0 \\ & \text{diag}(R) \geq \epsilon \cdot \mathbf{1} \\ & R_{11} = 1, \end{aligned} \quad (11)$$

where all non-diagonal elements of  $Q$  and  $R$  are set to zero. Both conditions  $Q \succeq 0$  and  $R \succeq \epsilon \cdot I_m$  dissolve naturally in linear inequality constraints for diagonal cost matrices.

**Remark 5** (Extension to block-diagonal cost matrices). *It is similarly possible to formulate an LP for block-diagonal  $Q$  and  $R$  if the blocks are of dimension two, i.e.*

$$Q = \begin{bmatrix} Q^1 & 0 & \cdots & 0 \\ 0 & Q^2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & Q^{k_q} \end{bmatrix}, \quad R = \begin{bmatrix} R^1 & 0 & \cdots & 0 \\ 0 & R^2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & R^{k_r} \end{bmatrix}$$

with  $Q^i, R^i \in \mathbb{R}^{2 \times 2}$ . Note that  $Q^{k_q}, R^{k_r}$  can also be scalar depending on the dimension of  $Q$  and  $R$ .

The additional constraints to impose  $Q \succeq 0$  and  $R \succ 0$  are obtained by applying Sylvester's criterion for positive definite matrices [18]:  $R$  is positive definite if  $R^i \succ 0$  for all  $i \in \{1, \dots, k_r\}$ , i.e.

$$\begin{aligned} r_{11}^i &\geq \epsilon \\ r_{22}^i &\geq \epsilon \\ r_{11}^i &\geq r_{12}^i + \epsilon \\ r_{12}^i &\geq -r_{11}^i + \epsilon \\ r_{22}^i &\geq r_{12}^i + \epsilon \\ r_{12}^i &\geq -r_{22}^i + \epsilon \end{aligned} \Leftrightarrow \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}}_{=:SR^i} \underbrace{\begin{bmatrix} r_{11}^i \\ r_{12}^i \\ r_{21}^i \\ r_{22}^i \end{bmatrix}}_{\text{vec}(R^i)} \geq \epsilon \cdot \mathbf{1}$$

for all  $i = 1, \dots, k_r$ . The condition  $R \succ 0$  results in

$$\underbrace{\begin{bmatrix} S^{R^1} & 0 & \cdots & 0 \\ 0 & S^{R^2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & S^{R^{k_r}} \end{bmatrix}}_{=:S^R} \begin{bmatrix} \text{vec}(R^1) \\ \text{vec}(R^2) \\ \vdots \\ \text{vec}(R^{k_r}) \end{bmatrix} \geq \epsilon \cdot \mathbf{1} \quad (12)$$

and  $R = R^\top$ . Note that  $S^{R^{k_r}} = 1$  if  $R^{k_r}$  is scalar.  $S^Q$  is determined analogously, where  $\epsilon = 0$  in (12) is sufficient to ensure that  $Q$  is positive semidefinite.

## V. PROPERTIES OF INVERSE OPTIMAL CONTROL SOLUTIONS

In this section, we present an explicit algebraic condition for inverse optimal control problems, which allows for deriving and analyzing the solution properties, such as uniqueness. First, we derive an algebraic solution for general cost function matrices. This result is then used to show that the inverse problem has a unique solution if the generating cost function matrices are diagonal. The derivations utilize the Kronecker product [16] to write (8) and (9) in vector form, where we define the vector variables

$$r := \text{vec}(R), \quad q := \text{vec}(Q), \quad y := \text{vec}(Y)$$

with  $r \in \mathbb{R}^{m^2}$ ,  $q \in \mathbb{R}^{n^2}$ , and  $y \in \mathbb{R}^{n(n-m)}$ . Using the reformulations in Section III-B, (8) and (9) are converted into matrix and vector notation resulting in the following system of equations

$$\begin{aligned} &\begin{bmatrix} (I_n \otimes I_n) \text{vec}(\Lambda) - ((A+BK)^\top \otimes A^\top) \text{vec}(\Lambda) \\ \text{vec}(\Lambda_p) \end{bmatrix} \\ &= \begin{bmatrix} M^R & M^Q & M^Y \\ N^R & 0 & N^Y \end{bmatrix} \begin{bmatrix} r \\ q \\ y \end{bmatrix} \end{aligned} \quad (13)$$

where  $M^R, N^R \in \mathbb{R}^{n^2 \times m^2}$ ,  $M^Q \in \mathbb{R}^{n^2 \times n^2}$ , and  $M^Y, N^Y \in \mathbb{R}^{n^2 \times n(n-m)}$  depend on  $A, B$ , and  $K$  and are defined in Appendix A.

### A. Algebraic condition for general cost matrices

In general, the generating cost matrices  $Q$  and  $R$  are not unique for a given control law solving the unconstrained optimal control problem with linear system dynamics (1), which is shown in the following by deriving the set of all solutions of the inverse optimal control problem. Let  $\Lambda = \Lambda_p = 0$  in (13), i.e.  $K$  is the solution of (1). Given an optimal feedback gain  $K$ , constraints (8) and (9) restrict the solutions  $Q, R$ , and  $Y$  to lie on a hyperplane, i.e.

$$\begin{bmatrix} r \\ q \\ y \end{bmatrix} \in \ker \left( \begin{bmatrix} M^R & M^Q & M^Y \\ N^R & 0 & N^Y \end{bmatrix} \right)$$

with convex boundaries reflecting  $Q \succeq 0$  and  $R \succ 0$ . This follows from the vectorized notation (13) and  $\Lambda = \Lambda_p = 0$ .

### B. Explicit algebraic solution for diagonal cost matrices

Using (13), we state sufficient conditions under which the cost function can be uniquely inferred from an optimal controller gain  $K$  if the generating cost matrices  $Q$  and  $R$  are diagonal. This result then directly leads to an explicit algebraic solution of the inverse problem.

Let  $i_Q \subset \{1, \dots, n^2\}$  be such that  $q_{i_Q} = \text{diag}(Q)$  contains all diagonal elements of  $Q$ , i.e.  $i_Q = \{(k-1)n + k \text{ for } k = 1, \dots, n\}$ , and let  $q_{\{1, \dots, n^2\} \setminus i_Q} = 0$ . The indices  $i_R \subset \{1, \dots, m^2\}$  are defined analogously. Then (13) with  $\Lambda = \Lambda_p = 0$  and  $r_1 = 1$  becomes

$$\begin{bmatrix} M_{*i_R \setminus 1}^R & M_{*i_Q}^Q & M^Y \\ N_{*i_R \setminus 1}^R & 0 & N^Y \end{bmatrix} \begin{bmatrix} r_{i_R \setminus 1} \\ q_{i_Q} \\ y \end{bmatrix} = - \begin{bmatrix} M_{*1}^R \\ N_{*1}^R \end{bmatrix}, \quad (14)$$

where  $r_{i_R \setminus 1} \in \mathbb{R}^{m-1}$ ,  $q_{i_Q} \in \mathbb{R}^n$ , and  $y \in \mathbb{R}^{n(n-m)}$ . Define

$$M := \begin{bmatrix} M_{*i_R \setminus 1}^R & M_{*i_Q}^Q & M^Y \end{bmatrix}, \quad N := \begin{bmatrix} N_{*i_R \setminus 1}^R & 0 & N^Y \end{bmatrix}.$$

The number of variables in (14) is

$$m-1 + n + n(n-m) = n^2 - (n-1)(m-1),$$

while  $M, N \in \mathbb{R}^{n^2 \times n^2 - (n-1)(m-1)}$ .

**Proposition 1** (Uniqueness of diagonal cost matrices). *Let  $K$  be the solution of (1) generated by diagonal cost matrices  $Q \succeq 0$  and  $R \succ 0$ . If*

$$\text{colrank} \left( \begin{bmatrix} M \\ N \end{bmatrix} \right) = m-1 + n + n(n-m), \quad (15)$$

then (14) with  $R_{11} = r_1 = 1$  has a unique solution and thereby also (5) has a unique solution for  $Q, R$  and  $P$ .

If (15) holds, then

$$\det \left( \begin{bmatrix} M \\ N \end{bmatrix}^\top \begin{bmatrix} M \\ N \end{bmatrix} \right) \neq 0$$

and the unique solution to the IOC problem is thus given by

$$\begin{bmatrix} r_{i_R \setminus 1}^* \\ q_{i_Q}^* \\ y^* \end{bmatrix} = - \left( \begin{bmatrix} M \\ N \end{bmatrix}^\top \begin{bmatrix} M \\ N \end{bmatrix} \right)^{-1} \begin{bmatrix} M \\ N \end{bmatrix}^\top \begin{bmatrix} M^R \\ N^R \end{bmatrix} \quad (16)$$

and  $r_1^* = 1$ . Eq. (16) provides an explicit algebraic solution for inverse optimal control problems if cost function matrices are diagonal and the measured  $K$  is an optimal gain.

**Remark 6.** *The solution of the LP in (11) and (16) are identical if the generating cost function matrices are diagonal and the measured  $K$  is an optimal gain. For non-optimal measured  $K$ , (16) can still be applied, however, different from (11), the solution may not be optimal for the considered cost in (11) and  $Q$  and  $R$  may not be positive (semi)-definite.*

### C. Feasibility of inverse optimal control problems

This section shows feasibility of the proposed inverse optimal control problems for optimal and non-optimal gains  $K$ . First, Theorem 1 proves feasibility of the LP formulation with diagonal cost matrices in (11). Then, Theorem 2 shows feasibility of the SDP in (10) using Theorem 1. Note that the algebraic condition in (16) is feasible if (15) holds.

**Theorem 1.** *The LP (11) is feasible for any  $K \in \mathbb{R}^{m \times n}$ .*

*Proof.* The proof can be found in the Appendix.  $\square$

**Theorem 2.** *The SDP (10) is feasible for any  $K \in \mathbb{R}^{m \times n}$ .*

*Proof.* Feasibility is implied by Theorem 1 as diagonal cost matrices also provide an admissible solution for (10).  $\square$

## VI. SIMULATION EXAMPLE

We analyze the properties of the different IOC formulations using the example system

$$x(k+1) = \begin{bmatrix} .6 & -.1 & .3 & .2 \\ 1 & 1 & 0 & 0 \\ .3 & .1 & .6 & -.3 \\ 0 & 0 & 1 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 1 & .2 \\ .5 & .4 \\ .2 & 1 \\ .5 & .5 \end{bmatrix} u(k).$$

The baseline optimal closed-loop gain  $K^{bl} \in \mathbb{R}^{2 \times 4}$  is defined as the solution of (1) with

$$Q^{bl} = \text{diag}([2 \ 3 \ 2 \ 1]), \quad R^{bl} = \text{diag}([2 \ 3]) \quad (17)$$

and is augmented with noise  $\nu$ , such that the measured gain matrix is defined by  $K_{ij} = K_{ij}^{bl}(1 + \nu_{ij})$  for all  $i, j$ , where  $\nu_{ij}$  is normally distributed with zero mean and standard deviation  $\sigma(\nu_{ij}) \in \{0, 0.01, \dots, 0.1\}$ . As performance measure for the proposed techniques we use the normalized deviation of the estimated from the optimal infinite-horizon cost, i.e.

$$\mathcal{E} := \left| \frac{\|P^*\|_2 - \|P^{cl}\|_2}{\|P^{cl}\|_2} \right|, \quad (18)$$

which offers a notion of closed-loop performance. We denote by  $P^*$ ,  $Q^*$ , and  $R^*$  the solution to the corresponding IOC problem.  $P^{cl}$  defines the infinite-horizon cost obtained from (5), i.e. the ARE with the solutions  $Q^*$  and  $R^*$ . In addition, we investigate the solver time to compute the objective function matrices.

Figure 1 shows the deviation  $\mathcal{E}$  in (18) for the four proposed formulations, i.e. the SDP in Section IV-B, the LP with block-diagonal cost matrices (LP<sub>b</sub>), cf. Remark 5, the LP with diagonal cost matrices (LP<sub>d</sub>) in Section IV-C, and the algebraic solution strategy (ALG) in (16) for varying standard deviation  $\sigma(\nu)$  and  $c_p = 1$  for all problems. The markers indicate the median of 1000 samples for  $\sigma(\nu) \in \{0, 0.01, \dots, 0.1\}$ , while the error bars indicate the 16<sup>th</sup> and 84<sup>th</sup> percentiles. For noise-free measurements, i.e.  $\sigma(\nu) = 0$ , all approaches recover the optimal cost function matrices in (17) such that  $\mathcal{E} = 0$ . The SDP is robust to noise with median value  $\mathcal{E} = 0$  for  $\sigma(\nu) \leq 0.1$ . LP<sub>b</sub> is more sensitive to noise but provides reliable solutions with median  $\mathcal{E} < 0.04$  for  $\sigma(\nu) \leq 0.1$ . LP<sub>d</sub> and ALG are more prone to noise with median values around  $\mathcal{E} = 0.06$  for  $\sigma(\nu) = 0.1$ .

Table I provides the computation times to solve the IOC problems, where the SDP is solved with MOSEK [19], LP<sub>b</sub> and LP<sub>d</sub> are solved with CPLEX [17], and the algebraic solution is computed in MATLAB. The hardware configuration is: 3.1 GHz Intel Core i7, 16 GB 1867 MHz DDR3, and Intel Iris Graphics 6100 1536 MB. The online computation capabilities of the convex IOC solutions are highlighted with solver times less than 3 ms for all proposed methods.

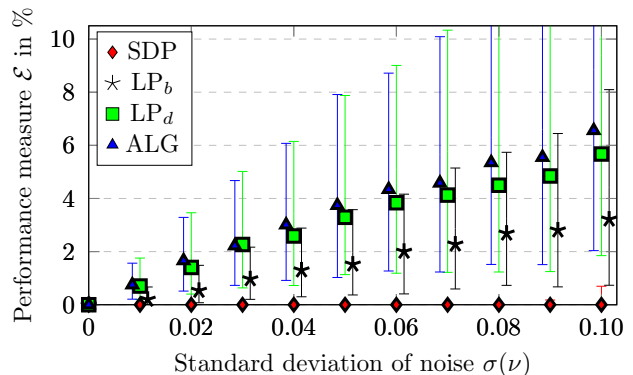


Fig. 1. Median and the 16<sup>th</sup> and 84<sup>th</sup> percentiles of the deviation  $\mathcal{E}$  from optimality for 1000 samples for SDP (red diamonds), LP<sub>b</sub> (black stars), LP<sub>d</sub> (green rectangles), and ALG (blue triangles).

TABLE I

SOLVER TIMES OF CONVEX INVERSE OPTIMAL CONTROL SOLUTIONS

Approach	Solver	Solver time
SDP	MOSEK	3.00 ms
LP <sub>b</sub>	CPLEX	0.565 ms
LP <sub>d</sub>	CPLEX	0.374 ms
ALG	MATLAB	< 0.1 ms

## VII. CONCLUSION

This paper showed that the inverse optimal control problem for recovering cost matrices of a quadratic cost function from a given feedback gain can be formulated as a convex optimization problem for both optimal and non-optimal closed-loop gains. Further, sufficient conditions were stated

under which this solution is unique providing an explicit algebraic expression for inferring diagonal cost function matrices. Simulation results have highlighted the robustness properties of the different formulations with respect to measured noise on the gain matrices and have demonstrated that the problems can be solved efficiently.

#### APPENDIX A: REFORMULATIONS

The matrices in (13) are defined as

$$M^R = \left( (K(A+BK)^{-1})^\top \otimes B^+ \right) - (K^\top \otimes A^\top B^+)$$

$$M^Q = I_n \otimes I_n = I_{n^2}$$

$$M^Y = ((A+BK)^\top \otimes A^\top B^{\perp\top}) - (I_n \otimes B^{\perp\top})$$

and

$$N^R = - \left( (K(A+BK)^{-1})^\top \otimes B^+ \right) + \left( B^+ \otimes (K(A+BK)^{-1})^\top \right)$$

$$N^Y = (I_n \otimes B^{\perp\top}) - (B^{\perp\top} \otimes I_n) T$$

with the identity matrix  $I_n$  and the permutation  $T$ , cf. (3).

#### APPENDIX B: LP FEASIBILITY PROOF

*Proof of Theorem 1.* We define

$$C := |(I_n \otimes I_n) - ((A+BK)^\top \otimes A^\top)|, \quad (19)$$

cf. (13), where we define the absolute value of a matrix as taking the absolute values of its elements:  $X = |Z| \Leftrightarrow X_{ij} = |Z_{ij}|$ . We introduce  $\lambda$ ,  $\lambda_p \in \mathbb{R}^{n^2 \times 1}$  as vector slack variables for  $\Lambda$ ,  $\Lambda_p$  in (13). With  $r_1 = 1$ , the LP (11) is equivalent to

$$\begin{aligned} & \min \|\lambda\|_1 + c_p \cdot \|\lambda_p\|_1 \\ \text{s.t.} \quad & \left[ \begin{array}{ccc} M_{*i_R \setminus 1}^R & M_{*i_Q}^Q & M^Y \\ N_{*i_R \setminus 1}^R & 0 & N^Y \end{array} \right] \begin{bmatrix} r_{i_R \setminus 1} \\ q_{i_Q} \\ y \end{bmatrix} + \begin{bmatrix} M_{*1}^R \\ N_{*1}^R \end{bmatrix} \leq \begin{bmatrix} C\lambda \\ \lambda_p \end{bmatrix} \\ & q_{i_Q} \geq 0 \\ & r_{i_R \setminus 1} \geq \epsilon \cdot \mathbf{1}, \end{aligned}$$

which can be cast as an LP in standard form, i.e.

$$\begin{aligned} & \min \|\lambda\|_1 + c_p \cdot \|\lambda_p\|_1 \\ \text{s.t.} \quad & Fw \leq f \end{aligned} \quad (20)$$

with

$$F = \begin{bmatrix} M_{*i_R \setminus 1}^R & M_{*i_Q}^Q & M^Y & -C & 0 & 0 \\ -M_{*i_R \setminus 1}^R & -M_{*i_Q}^Q & -M^Y & -C & 0 & 0 \\ N_{*i_R \setminus 1}^R & 0 & N^Y & 0 & -I_{n^2} & 0 \\ -N_{*i_R \setminus 1}^R & 0 & -N^Y & 0 & -I_{n^2} & 0 \\ -I_{m-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & -I_n & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$w = \begin{bmatrix} r_{i_R \setminus 1} \\ q_{i_Q} \\ y \\ \lambda \\ \lambda_p \end{bmatrix}, \quad f = \begin{bmatrix} -M_{*1}^R \\ M_{*1}^R \\ -N_{*1}^R \\ N_{*1}^R \\ -\epsilon \cdot \mathbf{1} \\ 0 \end{bmatrix}.$$

Assume (20) is infeasible, i.e.  $\{w \mid Fw \leq f\} = \emptyset$ . Then, according to Farkas' Lemma [20], there exists  $v =$

$[v_1^\top \ v_2^\top \ v_3^\top \ v_4^\top \ v_5^\top \ v_6^\top]^\top$ , where  $v_1, v_2, v_3, v_4 \in \mathbb{R}^{n^2}$ ,  $v_5 \in \mathbb{R}^{m-1}$ , and  $v_6 \in \mathbb{R}^n$  with

$$v^\top F = 0, \quad v^\top f < 0, \quad v^\top \geq 0. \quad (21)$$

From  $v^\top F = 0$ , it follows that

$$\begin{aligned} & (v_1^\top - v_2^\top) M_{*i_R \setminus 1}^R + (v_3^\top - v_4^\top) N_{*i_R \setminus 1}^R - v_5^\top = 0 \\ & (v_1^\top - v_2^\top) M_{*i_Q}^Q - v_6^\top = 0 \\ & -v_1^\top C - v_2^\top C = 0 \\ & -v_3^\top - v_4^\top = 0. \end{aligned}$$

It is immediate that  $v_1 = v_2 = v_3 = v_4 = 0$  because  $v \geq 0$  and  $C$  as in (19). It follows that  $v_5 = v_6 = 0$ . Thus the only  $v$  fulfilling  $v^\top F = 0$  is  $v = 0$ . If, however,  $v = 0$ , then  $v^\top f = 0$ , which contradicts assumption (21). Hence  $\{w \mid Fw \leq f\} \neq \emptyset$ .  $\square$

#### REFERENCES

- [1] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA: Athena Scientific, 2012.
- [2] R. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.
- [3] A. Y. Ng and S. J. Russell, "Algorithms for inverse reinforcement learning," in *Proc. 17th Int. Conf. Machine Learning*, Stanford, CA, 2000, pp. 663–670.
- [4] B. D. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey, "Maximum entropy inverse reinforcement learning," in *Proc. 23rd AAAI Conf. Artificial Intelligence*, Chicago, IL, 2008, pp. 1433–1438.
- [5] S. Levine, Z. Popovic, and V. Koltun, "Nonlinear inverse reinforcement learning with Gaussian processes," in *Proc. Neural Information Processing Systems Conf.*, Granada, Spain, 2011, pp. 19–27.
- [6] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proc. 21st Int. Conf. Machine Learning*, Banff, Canada, 2004, p. 1.
- [7] —, "Exploration and apprenticeship learning in reinforcement learning," in *Proc. 22nd Int. Conf. Machine Learning*, New York, NY, 2005, pp. 1–8.
- [8] A. Coates, P. Abbeel, and A. Y. Ng, "Learning for control from multiple demonstrations," in *Proc. 25th Int. Conf. Machine Learning*, New York, NY, 2008, pp. 144–151.
- [9] M. C. Priess, J. Choi, and C. Radcliffe, "Determining human control intent using inverse LQR solutions," in *Proc. ASME Dyn. Syst. Control Conf.*, Palo Alto, CA, 2013.
- [10] M. C. Priess, R. Conway, J. Choi, J. M. Popovich, and C. Radcliffe, "Solutions to the inverse LQR problem with application to biological systems analysis," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 2, pp. 770–777, Mar. 2015.
- [11] A. Ramadan, J. Choi, and C. J. Radcliffe, "Inferring human subject motor control intent using inverse MPC," in *Proc. American Control Conf.*, 2016, pp. 5791–5796.
- [12] E. Kreindler and A. Jameson, "Optimality of linear control systems," *IEEE Trans. Autom. Control*, vol. 17, no. 3, pp. 349–351, 1972.
- [13] M. Zanon, S. Gros, and M. Diehl, "Indefinite linear MPC and approximated economic MPC for nonlinear systems," *J. of Process Control*, vol. 24, no. 8, pp. 1273–1281, 2014.
- [14] 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.
- [15] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
- [16] K. B. Petersen and M. S. Pedersen, "The Matrix Cookbook," Nov. 2012.
- [17] IBM ILOG CPLEX, "V12. 1: Users Manual for CPLEX," *International Business Machines Corporation*, 2009.
- [18] G. T. Gilbert, "Positive definite matrices and Sylvester's criterion," *The American Mathematical Monthly*, vol. 98, no. 1, pp. 44–46, 1991.
- [19] MOSEK ApS, "The MOSEK optimization software," *Online: http://www.mosek.com/*, 2010.
- [20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.