2016 American Control Conference (ACC)
Boston Marriott Copley Place
July 6-8, 2016. Boston, MA, USA

(Un)ambiguous Optimal Control

Andrew Lamperski*, Bolei Di*, Tyler Lekang*, Jinsun Liu†, and Ran Tian*
* Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA, 55455
alampers@umn.edu, dixxx047@umn.edu, leka0024@umn.edu, tianx253@umn.edu
† Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA, 48109
jinsunl@umich.edu

***Abstract*— The phenomenon of using control actions to communicate is called signaling. Signaling is often beneficial in decentralized control problems with imperfect communication between agents. Seemingly simple control problems with signaling are often mathematically challenging. Humans, however, lack high-speed communication channels and routinely employ signaling strategies during cooperative movements. This paper presents a computationally tractable two-player problem that models several salient features of signaling problems arising in both decentralized control and human experiments. The problem consists of a signaler that reaches towards one of two possible targets, and an observer that decides on the target location based on noisy measurements of the movement. The signaler trades off control costs, such as energy, with informativeness for the observer. Two variants, the unambiguous case and the ambiguous case, are presented. In the unambiguous case, the signaler makes movements that are easy to distinguish, while in the ambiguous case, the signaler maximizes similarity of the movements. An approximation method for nonlinear systems is presented. When applied to a three-link arm model, the control scheme reproduces qualitative signaling phenomena observed in human reaching experiments.**

I. Introduction

Signaling phenomena, in which agents communicate with control actions, arise in decentralized control problems with imperfect communication channels. Signaling has been studied theoretically in decentralized control since the 1960s. Experimentally, signaling has been demonstrated in numerous cooperative movement experiments. This paper introduces a problem that is intended to form a bridge between the theoretical studies of signaling and phenomena observed in biological experiments.

A. Related Work

In decentralized control, the earliest example of signaling phenomena iss Witsenhausen's counterexample [1]. In this problem, one agent utilizes control actions to improve a partner's state estimate of a plant. To date, an optimal solution to Witsenhausen's problem is unknown and subsequent work has shown that problems with signaling can be computationally intractable [2], [3]. Recent work on signaling has identified control problems with signaling that can be solved computationally [4], [5]. See [6] for a modern treatment of signaling in control.

Humans lack perfect communication channels and so signaling with control actions would is useful for cooperation. Indeed, signaling phenomena, termed "coordination smoothers", have been observed in numerous human cooperation experiments [7]. Relevant to this paper, during cooperative target reaching tasks, humans exaggerate their movements in order to help their partners infer the location of the reaching target [8], [9].

Work on human-robot interaction has utilized automatic gesture recognition for decades [10], [11]. Recent results in robotics, [12], [13], approach similarly motivated problems as this paper and obtain qualitatively similar results, though the modeling and optimization techniques are different. Other recent studies have sought to enhance robotic cooperation by incorporating estimates of human intentions from movement data [14], [15]. Furthermore, in turn-taking tasks, such as hand-overs, human signaling phenomena have been exploited to improve control timing [16], [17]

B. Organization

The paper is organized as follows. The main coordination problem, called the *unambiguous linear quadratic regulator* is defined in Section II, and its solution is given in Section III. A related problem, called the *ambiguous linear quadratic regulator* is described in Subsection IV-A, anIn Subsection IV-B, the methods are extended to nonlinear systems. The method is applied to a three-link arm model, and qualitative features from a cooperative reaching experiment in [8] are reproduced. Conclusions are given in Section V.

II. The Unambiguous Linear Quadratic Regulator

A. Notation

For a sequence of vectors $x_0, x_1, \dots, x_N$, define $x_{0:N}$ by

$$x_{0:N} = \begin{bmatrix} x_0^\top & x_1^\top & \cdots & x_N^\top \end{bmatrix}^\top.$$

The probability of event A is denoted by $\mathbb{P}(A)$. The expected value of a random variable, X is denoted by $\mathbb{E}[X]$. The indicator function is denoted by χ.

The Euclidean norm is denoted by $\|\cdot\|$, so that $\|x\|^2 = x^\top x$. The determinant of a matrix M is denoted by $|M|$. The notation $M \succeq 0$ denotes that M is positive semidefinite, while $M \succ 0$ denotes that M is positive definite. The image of matrix M is denoted by $\mathrm{im}(M)$. The pseudo-inverse of a matrix is denoted by M^+.

B. A Two-Player Control and Decision Problem

This paper focuses on a two-player control and decision problem. The problem is designed to model several of the salient features from decentralized control and human cooperation experiments, discussed in Section I-A, while remaining mathematically tractable. The sequence of the problem is described as follows (see also Fig. 1):

1) The *signaler* begins at initial state $\hat{\mathbf{x}}$.
2) Based on the value of a binary random variable $b \in \{0, 1\}$, drawn uniformly, the signaler moves towards one of two possible targets τ^0 or τ^1. This results in a state trajectory, $x^0_{0:N}$ or $x^1_{0:N}$, depending on the value of $b \in \{0, 1\}$. We will assume that trajectories are generated by affine dynamics x^b_k, subject to a quadratic cost bound.
3) The *observer*, who is not given the value of b, collects noise-corrupted measurements of the movement:

$$z = y^b + w, \text{ where} \tag{1a}$$

$$y^b = \begin{bmatrix} C_0 x^b_0 \\ C_1 x^b_1 \\ \vdots \\ C_N x^b_N \end{bmatrix}, \quad w = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_N \end{bmatrix}. \tag{1b}$$

Here w_i are independent Gaussian random variables with mean 0 and covariance Σ_i.
4) The observer decides which target, τ^0 or τ^1, the signaler was reaching, based on the noisy measurements. The decision is denoted by $f(z)$. The observer's goal is to choose f so that it maximizes the probability of making the correct decision, $\mathbb{P}(f(z) = b)$.

Remark 1: It is assumed that the signaler and observer both know the possible strategies that the signaler can generate, $x^0_{0:N}$ and $x^1_{0:N}$. However, since the observer does not know b, it only measures the trajectories noisily, it does not know which trajectory the signaler used.

To complete the description of the problem, we must explain how the trajectories, $x^0_{0:N}$ and $x^1_{0:N}$ are generated. For $b = 0, 1$, the movement follows affine dynamics:

$$x^b_{k+1} = A^b_k x^b_k + B^b_k u^b_k + g^b_k, \quad x^b_0 = \hat{\mathbf{x}}, \tag{2}$$

To ensure that the signaler makes "reasonable" movements, the trajectories must satisfy a quadratic cost bound:

$$\mathbb{E}\left[\sum_{k=0}^{N}\left(x^{b\,\mathsf{T}}_k Q^b_k x^b_k + 2h^{b\,\mathsf{T}}_k x^b_k + u^{b\,\mathsf{T}}_k R^b_k u^b_k + 2\ell^{b\,\mathsf{T}}_k u^b_k\right)\right] \le c. \tag{3}$$

Explanation of the dynamics and cost bound: In (2), the superscript b is used to denote which target τ^b the signaler is attempting to reach and should not be interpreted as an exponent. For the basic linear quadratic theory, there is little loss of generality in assuming that time invariant dynamics, which are independent of the bit value b. In Subsection IV-B, however, we extend to the methods to nonlinear systems with dynamics of the form

$$x^b_{k+1} = F(x^b_k, u^b_k), \qquad x^b_0 = \hat{\mathbf{x}}, \tag{4}$$

and apply the iterative approximation scheme from [18]. In this case, (4) is linearized around trajectories $(x^b_{0:N}, u^b_{0:N})$, resulting in approximate dynamics of the form (2). In particular, the resulting matrices will be time varying, and will depend on which trajectory, $(x^0_{0:N}, u^0_{0:N})$ or $(x^1_{0:N}, u^1_{0:N})$, was used for linearization. Similarly, the general form of the costs bounds in (3) is required for quadratic approximations of non-quadratic cost bounds.

For convexity, it will be assumed that R^b_k is positive definite and Q^b_k is positive semidefinite. It will be assumed that (3) is strictly feasible.

The control action of the signaler has two roles. The first role is to regulate movement, as specified in (3), while the second role is to increase the probability that the observer makes the correct decision. The secondary role of improving the observer's decision probability can be interpreted as using control actions to transmit a single bit, b. Mathematically, the joint control and decision problem is modeled by the following optimization problem, which we call the "unambiguous linear quadratic regulator":

$$\max_{f,u} \; \mathbb{P}(f(z) = b) \tag{5a}$$

$$\text{s.t. } x^i_{k+1} = A^i_k x^i_k + B^i_k u^i_k + g^i_k, \; i = 0, 1 \tag{5b}$$

$$x^i_0 = \hat{\mathbf{x}}, \; i = 0, 1 \tag{5c}$$

$$\mathbb{E}\Big[\textstyle\sum_{k=0}^{N}\Big(x^{b\,\mathsf{T}}_k Q^b_k x^b_l + 2h^{b\,\mathsf{T}}_k x^b_k + u^{b\,\mathsf{T}}_k R^b_k u^b_k + 2\ell^{b\,\mathsf{T}}_k u^b_k\Big)\Big] \le c \tag{5d}$$

The main result of this paper, Theorem 1 in Section III-A, enables the solution of (5) based on iteratively solving linear quadratic regulator problems.

Remark 2: The term "unambiguous linear quadratic regulator" is given because the signaler minimizes the ambiguity of its trajectories, subject to a quadratic cost constraint.

Example 1: Consider the following special case of (2):

$$x^b_{k+1} = x^b_k + 0.1 u^b_k, \quad x^b_0 = 0.$$

The targets are given $x = \pm 1$, so that $\tau^b = (-1)^b$. The constraint on the signaling strategy, (3), is given by

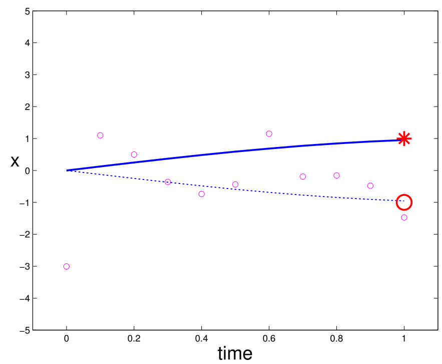
$$\mathbb{E}\left[\sum_{k=0}^{N} 0.01(u^b_k)^2 + (x^b_N - (-1)^b)^2\right] \le \tilde{c} \tag{6}$$

The left of (6) can be minimized using a linear quadratic regulator to give a minimal value of 0.182. Thus, in order for (6) to be strictly feasible, $\tilde{c}$ must be greater than 0.182.

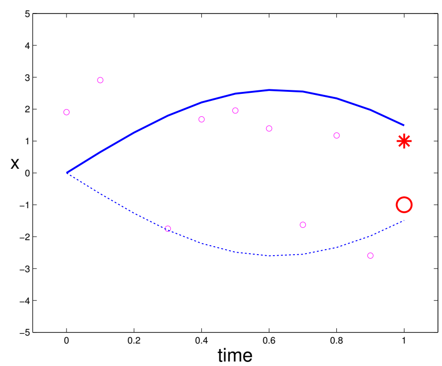
By varying the upper bound, $\tilde{c} > 0.182$, the trade-off between movement costs and observer accuracy can be changed. In particular, as $\tilde{c}$ increases, the trajectories become more exaggerated. See Fig. 1.

III. Solution

This section presents the main result in Subsection III-A, which shows that (5) problem is equivalent to a quadratic



(a) Low Signaling



(b) High Signaling

Fig. 1: Here, the signaler starts at $\hat{x} = 0$, and the targets are given by $x = \pm 1$. The solid line depicts the movement toward 1, while the dashed line depicts the movement toward the other target, -1. The magenta circles depict the noisy measurements which the observer uses to decide which target the signaler was trying to reach. Both cases correspond to optimal trajectories for the system from Example 1, but with different movement cost bound $\tilde{c}$ in (6). In 1a, $\tilde{c} = 0.2$ and the observer can correctly decide which target is being reached with probability 0.85. In 1b, $\tilde{c} = 4$, the observer can decide on the correct target with probability over 0.999.

maximin problem. Next, in Subsection III-B, a computational method for computing the optimal solution to (5) is given.

A. Main Result

Define Σ by $\Sigma = \mathrm{diag}(\Sigma_0, \Sigma_1, \dots, \Sigma_N)$. Then w is a zero-mean Gaussian random variable with covariance Σ.

Say that z has dimension m. If the control strategies of the signaler, $u^i_{0:N}$, are fixed, then $p(z, b)$ has a joint density given by

$$p(z,b) = p(b)p(z|b) \tag{7a}$$

$$= \frac{1}{2}\frac{1}{(2\pi)^{\frac{m}{2}}|\Sigma|^{\frac{1}{2}}}\int_{\mathbb{R}^m} \exp\left(-\frac{1}{2}(z-y^b)^\mathsf{T}\Sigma^{-1}(z-y^b)\right)dz. \tag{7b}$$

For compact notation, set $u = \begin{bmatrix} u^0_{0:N} \\ u^1_{0:N}\end{bmatrix}$.

The following theorem is the main result in the paper. It gives a characterization of the optimal strategies of the signaler and the observer.

Theorem 1: *For any signaling strategy, the optimal observer strategy is given by maximum likelihood decisions:*

$$f^*(z) = \begin{cases} 0 & \text{if } p(z|0) \geq p(z|1) \\ 1 & \text{if } p(z|1) > p(z|0). \end{cases} \tag{8}$$

A signaler strategy, u^*, is optimal if and only if it is optimal for the following maximin problem:

$$\max_{\lambda\geq 0}\min_u \begin{array}{l} -\sum_{k=0}^N (x^1_k - x^0_k)^\mathsf{T} C_k^\mathsf{T}\Sigma_k^{-1}C_k(x^1_k - x^0_k) \\ +\frac{\lambda}{2}\sum_{i=0}^1\sum_{k=0}^N \Big({x^i_k}^\mathsf{T} Q^i_k x^i_l + 2{h^i_k}^\mathsf{T} x^i_k \\ \quad +{u^i_k}^\mathsf{T} R^i_k u^i_k + 2{\ell^i_k}^\mathsf{T} u^i_k\Big) - \lambda c. \end{array} \tag{9a}$$

$$\text{s.t. } x^i_{k+1} = A^i_k x^i_k + B^i_k u^i_k + g^i_k,\ i = 0,1 \tag{9b}$$

$$x^i_0 = \hat{\mathbf{x}},\ i = 0,1 \tag{9c}$$

Remark 3: Note that the maximum likelihood decision strategy (8) can only be implemented if the signaler can evaluate the conditional distribution $p(z|b)$. In order to evaluate $p(z|b)$, the noise-free measurement vectors y^0 and y^1 must be known to the observer.

Proof of Theorem 1: Given data z, the observer's goal of deciding the value of the transmitted bit, b is a binary hypothesis testing problem. When $p(b = 0) = p(b = 1) = \frac{1}{2}$, a classical result on hypothesis testing, [19], shows that the minimum error probability is given by (8).

Given any y^0 and y^1 vectors, an argument from Gaussian detection theory, [20], shows that the probability that the observer, using (8), makes the correct decision is given by:

$$\mathbb{P}(f^*(z) = b) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\Delta} e^{-\frac{t^2}{2}}dt, \tag{10}$$

where $\Delta \geq 0$ is defined by

$$\begin{aligned}(2\Delta)^2 &= \left(y^1 - y^0\right)^\mathsf{T}\Sigma^{-1}\left(y^1 - y^0\right) \\ &= \sum_{k=0}^N (x^1_k - x^0_k)^\mathsf{T} C_k^\mathsf{T}\Sigma_k^{-1}C_k(x^1_k - x^0_k).\end{aligned} \tag{11}$$

It follows that maximizing the probability of correctness is equivalent to maximizing $(2\Delta)^2$. Since the equality constraints, (5b) and (5c), are affine, x^i_k can be computed as affine functions of the input u and the fixed parameters $\hat{\mathbf{x}}$ and g^i_k. Let $x^i_k(u)$ denote the resulting affine functions of u. Eliminating the equality constraints, combined with (11), implies that (9) is equivalent to the following problem:

$$\min_u\ -\sum_{k=0}^N (x^1_k(u) - x^0_k(u))^\mathsf{T} C_k^\mathsf{T}\Sigma_k^{-1}C_k(x^1_k(u) - x^0_k(u)) \tag{12a}$$

$$\text{s.t. } \frac{1}{2}\sum_{i=0}^1\sum_{k=0}^N \Big({x^i_k(u)}^\mathsf{T} Q^i_k x^i_l(u) + 2{h^i_k}^\mathsf{T} x^i_k(u) \tag{12b}$$

$$+{u^i_k}^\mathsf{T} R^i_k u^i_k + 2{\ell^i_k}^\mathsf{T} u^i_k\Big) \leq c.$$

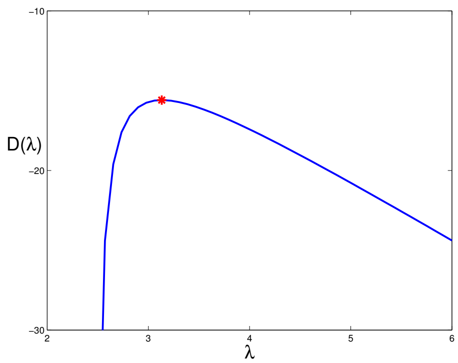


Fig. 2: The dual function, $D(\lambda)$, corresponding to the system from Example 1, with $\tilde{c} = 4$. Here, the optimal dual solution λ^* is strictly feasible, and so Lemma 1 in the Appendix implies that the optimal signaler strategy is unique.

The reformulated quadratic cost constraint, (12b), holds because we assumed that $\mathbb{P}(b = 0) = \mathbb{P}(b = 1) = \frac{1}{2}$.

The assumptions $R_k^i \succ 0$ and $Q_k^i \succeq 0$ imply that (12) is a quadratic optimization with a single, strictly convex quadratic constraint. Thus, (12) is equivalent to a problem of the form

$$\min_u \; u^\top P_0 u + 2b_0^\top u + c_0 \tag{13a}$$

$$\text{s.t.} \;\; u^\top P_1 u + 2b_1^\top u + c_1 \leq 0, \tag{13b}$$

where P_1 is positive definite. Since it is assumed that (3) is strictly feasible, the reformulated constraint (13b) is assumed to be strictly feasible as well.

The Lagrangian of (13) is

$$L(u, \lambda) = u^\top (P_0 + \lambda P_1) u + 2(b_0 + \lambda b_1)^\top u + c_0 + \lambda c_1 \tag{14}$$

and the corresponding dual problem is

$$\max_{\lambda \geq 0} \; - (b_0 + \lambda b_1)^\top (P_0 + \lambda P_1)^+ (b_0 + \lambda b_1) + c_0 + \lambda c_1 \tag{15a}$$

$$\text{s.t.} \;\; P_0 + \lambda P_1 \succeq 0 \tag{15b}$$

$$(b_0 + \lambda b_1) \in \operatorname{im}(P_0 + \lambda P_1) \tag{15c}$$

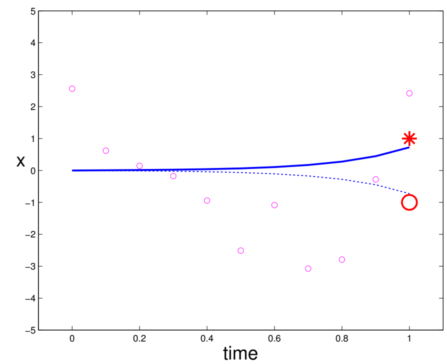
Note that the maximin problem from (9) is equivalent to the maximin problem used to compute the Lagrange dual:

$$\max_{\lambda \geq 0} \min_u L(u, \lambda) = \max_{\lambda \geq 0} D(\lambda). \tag{16}$$

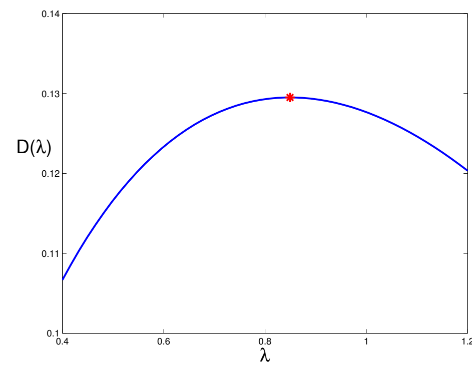
Lemma 1 in the Appendix implies that (15) is a strong dual of (13), and so u^* is optimal for (13) if and only if there is there is value $\lambda^* \geq 0$ such that (u^*, λ^*) are optimal for the maximin problem (16). The proof of the theorem is complete, since (16) is equivalent to the (9). ■

B. Computing the Optimal Signaling Strategy

Now, a procedure for solving the maximin problem in (9) will be described. For each value of $\lambda \geq 0$, the inner minimization of (9) is a linear quadratic regulator problem with optimal value given by $D(\lambda)$, the dual function from



(a) Signaler Strategies



(b) Dual Function

Fig. 3: 3a. This figure depicts the optimal signaling strategy for the problem defined in (17), as applied to the system from Example 1. In this case, $\tilde{c} = 0.4$, and the probability that the observer identifies the correct target is 0.68. 3b To find the optimal signaler strategy, the dual function, $D(\lambda)$, is maximized. In this case, the dual optimum occurs at $\lambda^* > 0$, so the optimal signaler strategy is unique.

(16). (Below some value, $\lambda < \lambda_l$, the regulator problem is unbounded below and so $D(\lambda) = -\infty$.)

Since $D(\lambda)$ is a Lagrangian dual function, it is a concave. Furthermore, since λ is a scalar, it can be efficiently maximized numerically. Furthermore, if a maximizing solution, λ^*, is found with $\lambda^* > \lambda_l$, then the signaling strategy, u^*, computed from the corresponding regulator problem is unique. See Fig. 2.

IV. Variations and Extensions

A. Ambiguous Optimal Control

So far, the signaler's objective has been to increase the observer's decision accuracy, while maintaining a control performance specified in (3). Say, instead, that the signaler intends to decrease decision accuracy. This scenario arises in game settings, such as when a player disguises the direction that they throw a ball. This scenario of "ambiguous optimal control" can be modeled by changing the objective in the original problem, (5), to a maximin problem:

$$\max_f \min_u \ \mathbb{P}(f(z) = b) \tag{17a}$$

$$\text{subject to} \quad \text{Constraints (5b), (5c), and (5d)} \tag{17b}$$

In this new formulation, the arguments from Subsection III-A apply, with the only change being that the objective from (12a) will have a positive sign instead of a negative sign. Thus, the following theorem holds.

Theorem 2: *An optimal observer strategy is given by maximum likelihood decisions,* (8).

A signaler strategy, u^*, is optimal for (17) *if and only if it is optimal for the following maximin problem:*

$$\max_{\lambda \geq 0} \min_u \begin{array}{l} \sum_{k=0}^N (x_k^1 - x_k^0)^\top C_k^\top \Sigma_k^{-1} C_k (x_k^1 - x_k^0) \\ +\frac{\lambda}{2} \sum_{i=0}^1 \sum_{k=0}^N \Big({x_k^i}^\top Q_k^i x_l^i + 2{h_k^i}^\top x_k^i \\ \quad + {u_k^i}^\top R_k^i u_k^i + 2{\ell_k^i}^\top u_k^i \Big) - \lambda c. \end{array} \tag{18a}$$

$$\text{s.t.} \ x_{k+1}^i = A_k^i x_k^i + B_k^i u_k^i + g_k, \ i = 0, 1 \tag{18b}$$

$$x_0^i = \hat{\mathbf{x}}, \ i = 0, 1. \tag{18c}$$

Note that for any $\lambda > 0$, the inner minimization in (18) is a strictly convex linear quadratic regulator problem. As discussed in Subsection III-A, the value of the inner minimization for fixed λ is the dual objective of a minimization problem, $D(\lambda)$. So, as before, the optimal signaler strategy can be computed by maximizing $D(\lambda)$, which is concave. If the optimal dual solution occurs with $\lambda^* > 0$, then the optimal signaler solution u^* is unique. See Fig. 3.

B. Nonlinear Systems

The theory can also be used to find approximate signaling strategies by applying Theorem 1 or 2 to a linear-quadratic approximation. Say that the nonlinear dynamics are given by

$$x_{k+1}^i = F(x_k^i, u_k^i), \quad x_0^i = \hat{\mathbf{x}}, \tag{19}$$

for $i = 0, 1$, and the trajectory constraint is given by

$$\frac{1}{2} \sum_{i=0}^1 \sum_{k=0}^N \left(q_k(x_k^i) + r_k(u_k^i) \right) \leq c. \tag{20}$$

Consider noisy measurements given by

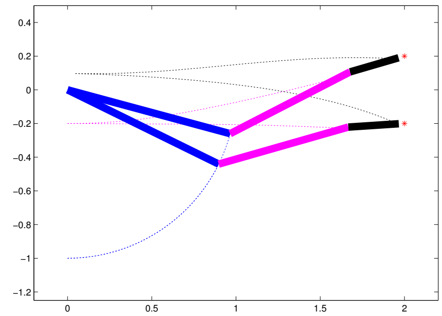
$$z_k = H_k(x_k^b) + w_k = y_k^b + w_k, \tag{21}$$

where w_k are independent zero mean Gaussians with covariance Σ_k.

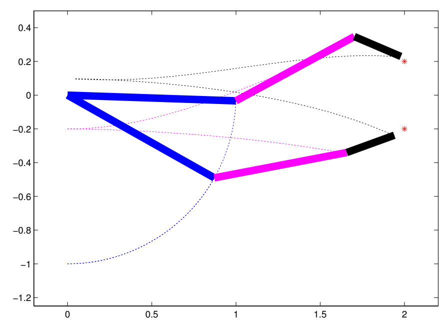
As in the linear case, if the observer knows y^0 and y^1, it has an optimal strategy based on maximum likelihood decisions, (8). Using arguments from detection theory, as in the linear case, it can be shown that $\mathbb{P}(f^*(z) = b)$ increases monotonically with

$$\sum_{k=0}^N (H_k(x_k^1) - H_k(x_k^0))^\top \Sigma_k^{-1} (H_k(x_k^1) - H_k(x_k^0)). \tag{22}$$

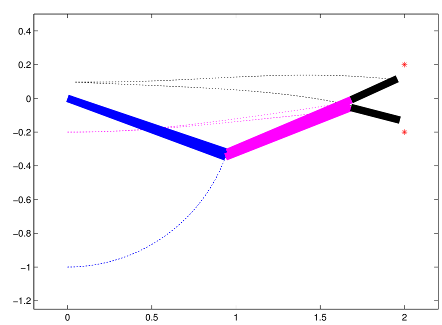
Approximately optimal solutions for the signaler can be computed by successively finding linear approximations to



(a) Locally Optimal Reaches



(b) Signaling



(c) Deception

Fig. 4: Trajectories of a three-link arm model. There are two possible targets, represented by red $*$ marks. The dotted lines depict the trajectories of the link tips. In the no signaling case, 4a, the trajectories are a local minimum of the movement cost, (20). In 4b, the signaler attempts to maximize the decision accuracy and the trajectories exhibit "wrist-pointing" phenomena observed in experiments. In 4c, the movements are deceptive.

the dynamics, (19) and quadratic approximations to (20) and (22), and then solving either (9) or (18). See [18].

Example 2: Fig. 4 shows the result of the algorithm as applied to reaching movements of three-link arm. This problem is designed to model qualitative features of an experiment from [8]. In this experiment, a leader and follower coordinate arm movements to grasp bars at either a high or low targets. Only the leader receives a cue about the desired target. In experiments, the leaders modulated the height of their wrists during movements to provide information to followers about the target. These features are captured by the differences in the third joint in Fig. 4

The states of the nonlinear system are given by joint angles and angular velocities: $x^\mathsf{T} = \begin{bmatrix}\theta_1 & \theta_2 & \theta_3 & \dot\theta_1 & \dot\theta_2 & \dot\theta_3\end{bmatrix}^\mathsf{T}$ For cleaner notation, the time indices k and target indices i will be dropped in the description of the system.

The observer noisily measures the link tip locations:

$$y_1 = \ell_1 \begin{bmatrix}\cos(\theta_1)\\ \sin(\theta_2)\end{bmatrix}, \quad y_2 = y_1 + \ell_2 \begin{bmatrix}\cos(\theta_1+\theta_2)\\ \sin(\theta_1+\theta_2)\end{bmatrix}$$

$$y_2 = y_3 + \ell_3 \begin{bmatrix}\cos(\theta_1+\theta_2+\theta_3)\\ \sin(\theta_1+\theta_2+\theta_3)\end{bmatrix}$$

Inputs are given by torques applied at the joints. The other external torques on the joints are due to gravity and viscous damping. Continuous-time equations of motion are computed by the Euler-Lagrange equations, [21]. Discrete-time equations of the form in (19) are computed using a first-order Euler approximation of the Euler-Lagrange equations.

The movement cost bound is of the form

$$\frac{1}{2}\left(\sum_{k=0}^{N} \alpha \|u_k^i\|^2 + \beta \|y_{N,3}^i - \tau^i\|^2\right) \le c,$$

where τ^i are spatial targets and $y_{N,3}^i$ are the Cartesian coordinates of the third link tip at the final time.

V. Conclusion

This paper presented a two-player problem in which one player balances control costs with signaling strength. This two-person problem was reduced to an equivalent maximin problem, which can be solved by iterating over linear quadratic regulator problems. An approximation scheme for nonlinear control problems was also sketched. When applied to an arm model, the method produced qualitative phenomena observed in human experiments.

References

[1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal of Control*, vol. 6, no. 1, 1968.

[2] J. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Transactions on Automatic Control*, vol. 30, no. 5, pp. 440–446, 1985.

[3] C. V. Goldman and S. Zilberstein, "Decentralized control of cooperative systems: Categorization and complexity analysis," *Journal of Artificial Intelligence Research*, vol. 22, pp. 143 – 174, 2004.

[4] C. Amato, G. Konidaris, G. Cruz, C. A. Maynor, J. P. How, and L. P. Kaelbling, "Planning for decentralized control of multiple robots under uncertainty," in *Proceedings of the Workshop on Planning and Robotics (PlanRob) at the Twenty-Fourth International Conference on Automated Planning and Scheduling (ICAPS-14)*. AAAI Press, 2014.

[5] Y. Ouyang and D. Teneketzis, "Balancing through signaling in decentralized routing," in *IEEE Conference on Decision and Control*, 2014, pp. 1675 – 1680.

[6] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Birkhäuser, 2013.

[7] C. Vesper, S. Butterfill, G. Knoblich, and N. Sebanz, "A minimal architecture for joint action," *Neural Networks*, vol. 23, pp. 998 – 1003, 2010.

[8] L. M. Sacheli, E. Tidoni, E. F. Pavone, S. M. Aglioti, and M. Candidi, "Kinematics fingerprints of leader and follower role-taking during cooperative joint actions," *Exp Brain Res*, vol. 226, pp. 473 – 486, 2013.

[9] C. Vesper and M. J. Richardson, "Strategic communication and behavioral coupling in asymmetric joint action," *Exp Brain Res*, vol. 232, pp. 2945 – 2956, 2014.

[10] S. Waldherr, R. Romero, and S. Thrun, "A gesture based interface for human-robot interaction," *Autonomous Robots*, vol. 9, pp. 151 – 173, 2000.

[11] N. Mavridis, "A review of verbal and non-verbal humanrobot interactive communication," *Robotics and Autonomous Systems*, vol. 63, pp. 22 – 35, 2015.

[12] A. Dragan, K. Lee, and S. Srinivasa, "Legibility and predictability of robot motion," in *Human-Robot Interaction*, March 2013.

[13] A. Dragan, R. Holladay , and S. Srinivasa, "Deceptive robot motion: Synthesis, analysis and experiments," *Autonomous Robots*, July 2015.

[14] Z. Wang, K. Mülling, M. P. Deisenroth, H. B. Amor, D. Vogt, B. Schölkopf, and J. Peters, "Probabilistic movement modeling for intention inference in human-robot interaction," *International Journal of Robotics Research*, vol. 32, no. 7, pp. 841 – 858, 2013.

[15] Y. Li and S. S. Ge, "Human-robot collaboration based on motion intention estimation," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 1007 – 1014, 2014.

[16] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli, "Toward seamless humanrobot handovers," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 112 – 132, 2013.

[17] A. Moon, D. M. Troniak, B. Gleeson, M. K. Pan, M. Zheng, B. A. Blumer, K. MacLean, and E. A. Croft, "Meet me where I'm gazing: How shared attention gaze affects human-robot handover timing," in *ACM/IEEE International Conference on Human-robot Interaction*, 2014, pp. 334–341.

[18] W. Li and E. Todorov, "Iterative linear quadratic regulator design for nonlinear biological movement systems," in *International Conference on Informatics in Control, Automation and Robotics*, 2004, pp. 222 – 229.

[19] B. C. Levy, *Principles of signal detection and parameter estimation*. Springer Science & Business Media, 2008.

[20] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[21] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Locomotion*. CRC Press, 1994.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

Appendix

The following lemma is required for the duality arguments of this paper. It specializes the results of Appendix B [22] to the case that $P_1 \succ 0$. The proof is omitted for space purposes.

Lemma 1: *The following hold*

1) *There is zero duality gap between* (13) *and* (15).
2) *The optimal value of* (13) *is finite.*
3) *If λ is dual feasible, then all $\hat\lambda > \lambda$ are also dual feasible, and the corresponding primal minimizer, $\hat u$, is uniquely defined by*

$$\arg\min_u L(\hat\lambda, u) = -(P_0 + \hat\lambda P_1)^{-1}(b_0 + \hat\lambda b_1) \quad (23)$$