Bio-Inspired Optimal Control via Intermittent Cooperation*

Cheng Shao[†] and D. Hristu-Varsakelis[‡]

Abstract— **We investigate the solution of a large class of fixed-final-state optimal control problems by a group of cooperating dynamical systems. We present a pursuit-based algorithm – inspired by the foraging behavior of ants – that requires each system-member of the group to solve a finite number of optimization problems as it follows other members of the group from a starting to a final state. Our algorithm, termed "sampled local pursuit", is iterative and leads the group to a locally optimal solution, starting from an initial feasible trajectory. The proposed algorithm is broad in its applicability and generalizes previous results; it requires only short-range sensing and limited interactions between group members, and avoids the need for a "global map" of the environment or manifold on which the group evolves. We include simulations that illustrate the performance of our algorithm.**

I. INTRODUCTION

In nature, many animal groups exhibit highly organized and efficient "collective behaviors", despite their members' limited intelligence. For instance, worker honey bees can coordinate their distribution among different flowers in accordance with the profitability of each source; a school of fish can move together in a tight formation; ants can recruit nest-mates to form efficient foraging trails [1], [2], [3]. These examples illustrate how aggregate behavior may be qualitatively different from individual actions and that cooperation among members of a natural collective helps them overcome their limitations and accomplish complex tasks that may be impossible for them to attain individually.

Observations of the qualitatively similar behaviors of members of animal groups, coupled with their cognitive and physical limitations, support the conclusion that their collective efficiency and elegance are self-organized and must be "encoded" in fairly simple patterns (as far as individual actions are concerned), in contrast to the complex performance of the group. Moreover, many of the tasks performed by natural groups are functionally similar to what one might require from engineered collectives. In some cases, members of a biological group and those of a decentralized group of autonomous systems operated under similar constraints in the sense that they are both usually equipped with limited sensing, communication and computing capabilities.

The potential of a group to "be more than the sum of its members" has already seeded a variety of recent research directions in system and control community, from modeling of animal groups [1], [4], [5], to distributed collective covering and searching [6], [7], estimating by groups [8], [9], biologically-motivated optimization [10], [11] and cooperative robotic teams [12], [13]. The objective of this paper is to investigate the biologically-inspired cooperative solution to a class of optimal control problems with fixed final states. We are particularly interested in applying models of the foraging behavior of ants, which are well-known path optimizers (see for example [1]).

One of the early optimization methods inspired by trail formation in ants was presented in [4], where it was shown that ants that "pursued" one another on $\mathbb{R}^2$ (each pointing its velocity vector towards a predecessor) had the effect of producing progressively "straighter" trails. That idea was later extended to kinematic vehicles moving on non-Euclidean environments [14]. Both these works were restricted exclusively to the "discovery" of geodesics, meaning that the autonomous systems-members of the group had very simple dynamics with no drift terms. In this paper we show that the earlier work can be generalized to a much broader class of optimal control problems, and agents[1] with non-trivial dynamics. We propose an iterative, decentralized algorithm that involves "local pursuit" (to use the term coined in [4]), of members of a collective, this time in a broader and more intricate setting. Our algorithm has lower computational requirements than previous "continuous pursuit" formulations [15], [16] and requires agents to communicate with their neighbors a finite number of times. The agents do not need a global map of their environment or even an agreed-upon common coordinate system. The proposed algorithm is most useful in trajectory optimization problems which are easier to solve when boundary conditions are "close" to one another (because of, for example, the agents' computational or sensing limitations), with the term "close" taken to include not only geographical separation but also distance on the manifold on which copies of a dynamical system evolve.

The remainder of this paper is organized as follows: In Sec. II we describe the class of optimal control problems we are concerned with and propose an iterative, decentralized solution, termed "sampled local pursuit", which is inspired by the foraging behavior of ants. Section III contains the main results discussing the behavior of a collection of control systems evolving under the proposed algorithm. Sec. IV presents a pair of simulations designed to illustrate the performance of sampled local pursuit.

* This work was supported by the NSF under Grant . EIA0088081 and by ARO ODDR&E MURI01 Grant No. DAAD19-01-1-0465 (Center for Communicating Networked Control Systems, through Boston University).

† Corresponding author; Department of Mechanical Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA, cshao@glue.umd.edu

‡ Department of Applied Informatics, University of Macedonia, Thessaloniki 54006 Greece, dcv@uom.gr

[1] Throughout the paper we will use "agent" to refer to a member of a group of control systems.

0-7803-9098-9/05/$25.00 ©2005 AACC

II. A DISCRETE BIO-INSPIRED ALGORITHM FOR OPTIMAL CONTROL

We are interested in optimal control problems using a group of cooperating "agents". For our purposes, each agent is a "copy" of a dynamical system:

$$\dot{x}_k = f(x_k, u_k), \qquad x_k(t) \in \mathbb{R}^n, u_k(t) \in \Omega \subset \mathbb{R}^m \tag{1}$$

for $k = 0, 1, 2 \ldots$. Physically, each copy of (1) could stand for a robot, UAV or other autonomous system. Each $x_k(t) : [0, T] \rightarrow \mathbb{R}^n$ represents a trajectory defined by the k^{th} agent's evolution.

A. Problem Statement and Notation

Assume that there is a pair of states x_0 and x_f which are equilibrium points[2] of (1) for $u = 0$. The problem we are concerned with is finding a trajectory $x^*(t) \ \ (t \in [0, T], T$ fixed) that minimizes

$$J(x, \dot{x}, t_0, T) = \int_{t_0}^{t_0+T} g(x(t), \dot{x}(t), t)dt \tag{2}$$

with $x(t_0) = x_0, \ x(t_0{+}T) = x_f, \ g(\cdot, \cdot, \cdot) \geq 0$, and subject to (1).

It will be convenient to define the following notation. Let $\mathbb{D} \subset \mathbb{R}^n$ be a domain containing states a and b. Assume $0 < \sigma \leq T$ and $t_0 \geq 0$. The optimal trajectory from a to b in fixed T units of time will be denoted by $x^*(t) \ \ (t \in [t_0, t_0 + T])$ satisfying:

$$J(x^*, \dot{x}^*, t_0, T) = \min_x J(x, \dot{x}, t_0, T), \tag{3}$$

subject to $x(t_0) = a, \ x(t_0 + T) = b$. We define the cost of following the optimal trajectory from a to b for σ units of time with:

$$\eta(a, b, T, t_0, \sigma) \triangleq \int_{t_0}^{t_0+\sigma} g(x^*(t), \dot{x}^*(t), t)dt \qquad \sigma \leq T \tag{4}$$

where the optimal trajectory $x^*(t)$ is defined in (3).

For a generic trajectory $x(t)$ of (1), we define

$$C(x, t_0, \sigma) \triangleq \int_{t_0}^{t_0+\sigma} g(x(t), \dot{x}(t), t)dt \tag{5}$$

to be the cost incurred along $x(t)$ during $[t_0, t_0 + \sigma)$.

The following can be derived easily from the properties of optimal trajectories and are helpful in future argument.

Fact: Let η, C be defined by (4),(5), let $x_k(t)$ be a trajectory of (1) and $x^*(t)$ an optimal trajectory of (3). Then:

1) *$\eta(a, b, T, t_0, \sigma) \leq C(x_k, t_0, \sigma)$ for any $x^*(t)$ satisfying (3) with $x_k(t_0) = x^*(t_0), \ x_k(t_0 + \sigma) = x^*(t_0 + \sigma)$.*
2) *$\eta(a, c, T, t_0, T) \leq \eta(a, b, \sigma, t_0, \sigma) + \eta(b, c, T - \sigma, t_0 + \sigma, T - \sigma)$*
3) *$C(x_k, t_0, T) = C(x_k, t_0, \sigma) + C(x_k, t_0 + \sigma, T - \sigma)$*
4) *$\eta(a, b, T, t_0, \sigma) = C(x^*, t_0, \sigma)$.*

B. A Pursuit-based Optimal Control Algorithm

We assume that we have available an initial feasible (but sub-optimal) control/trajectory pair $(u_{feas}(t), x_{feas}(t))$ for (1), obtained through a combination of a-priori knowledge about the problem and/or random exploration. We consider the formation of an ordered sequence of agents, with each agent trying to reach its predecessor along an optimal trajectory. The sequence is initiated with the first agent following x_{feas} to the desired final state. The precise rules that govern the movement of each agent are:

Algorithm (Sampled Local Pursuit): *Identify two states x_0 and x_f on $\mathbb{D}$. Let $x_0(t) \ \ (t \in [0, T])$ be an initial trajectory satisfying (1) with $x_0(0) = x_0, x_0(T) = x_f$. Choose Δ, $\delta \in \mathbb{R}$ such that $0 < \delta < \Delta \leq T$. Then:*

1) *For $k = 1, 2, 3 \ldots$, let $t_k = k\Delta$ be the starting time of the k^{th} agent, i.e. $u_k(t) = 0, \ x_k(t) = x_0$ for $0 \leq t \leq t_k$.*
2) *When $t = t_k + i\delta, i = 0, 1, 2, 3, \ldots$, calculate the control $u^*_t(\tau)$ that achieves (subj. to (1)):*
$$\begin{cases} \eta(x_k(t), x_{k-1}(t), \Delta, t, \Delta), \ \tau \in [t, t + \Delta] \\ \qquad\qquad\qquad\qquad \text{if } \Delta + i\delta < T \\ \eta(x_k(t), x_f, \lambda, t, \lambda), \ \tau \in [t, t_k + T] \\ \qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$
where $\lambda = t_k + T - t$.
3) *Apply $u_k(t) = u^*_{t_k+i\delta}(t - t_k - i\delta)$ to the k^{th} agent for $t \in [t_k + i\delta, t_k + (i + 1)\delta)$ if $\Delta + i\delta < T$, or for $t \in [t_k + i\delta, t_k + T)$ otherwise.*
4) *Repeat from step 2 until the k^{th} agent reaches x_f.*

There are two adjustable parameters in the sampled local pursuit (SLP) algorithm: the "following interval" Δ which denotes the frequency with which new agents depart from x_0, and the "updating interval" δ which denotes the frequency with which an agent samples the state of his predecessor. To illustrate the pursuit process, we refer to the $(k-1)^{th}$ agent as the "leader" and to the k^{th} agent as the "follower". We will refer to the times $t^i_k = t_k + i\delta, i = 0, 1, 2, 3 \ldots$ as the "updating times". At every updating time, the follower finds an optimal trajectory from itself to its leader, and moves on it during $[t^i_k, t^i_k + \delta]$, until next updating time. This process continues until the follower reaches the final state. Usually we take $0 < \delta < \Delta$ so that each agent only needs to solve a finite number of optimal control problems. If the problem in question can be solved efficiently, one may choose to decrease δ. In fact, the case $\delta \rightarrow 0$ leads to a continuous version of the SLP algorithm (of which [14] and [4] are special cases), where each agent is constantly updating its trajectory in response to its leader's movement (see [16], [15]). The idea of optimizing a trajectory by applying a repeatedly-updated sequence of controls is also present in model predictive control (MPC). However, in MPC one typically computes the optimal control from the current to the terminal state x_f (see, for example, [17] and [18]), which in the present situation we would like to avoid. Instead, SLP substitutes

[2] Without loss of generality we assume that $u = 0$ at those equilibria.

a series of "shortened" versions of the original problem, involving only trajectories linking leader and follower. In the next section we show that SLP leads to (locally) optimal trajectories.

III. MAIN RESULTS

Recall that the proposed algorithm defines an ordered sequence of trajectories $\{x_k(t)\}$. We would like to investigate the properties of the limiting trajectory generated by the group, i.e. $x_k(t)$ as $k \to \infty$. We begin by discussing convergence of the iterated trajectories.

Lemma 1: (Convergence of Cost): Assume a group of agents $x_0, x_1, \ldots, x_k$ evolve under SLP with starting state x_0 and target state x_f. Suppose an initial control/trajectory pair, $\{u_0(t), x_0(t)\}$ $(t \in [0, T])$, satisfying $x_0(t) = x_0$ and $x_0(T) = x_f$ is given. If the updating time satisfies $0 < \delta \leq \Delta$, then the cost of the iterated trajectories will converge, i.e. $\lim_{k\to\infty} C(x_k, t_k, T)$ exists.

Proof: Consider the pursuing process between the $(k-1)^{th}$ and k^{th} agents. As shown in Fig. 1, the dotted line, denoted by $x_{k-1}(t)$ on $[t_{k-1}, t_{k-1}+T]$, indicates the leader's path. The solid lines, denoted by $x_k(t)$, are the realized trajectories of the "follower", and the dashed lines, noted by $\hat{x}_k(t)$, are the planned trajectories along which the follower plans to move at $t_k + i\delta$ but may not do so because it will update its future trajectory at $t_k+(i+1)\delta$. For $t \in [t_k, t_k+\delta]$, the follower moves on an optimal trajectory from state $x_k(t_k)$ to $x_{k-1}(t_k)$ over Δ units of time. Thus from Fact 1:

$$\eta(x_k(t_k), x_{k-1}(t_k), \Delta, t_k, \Delta) \quad \leq \quad C(x_{k-1}, t_{k-1}, \Delta)$$

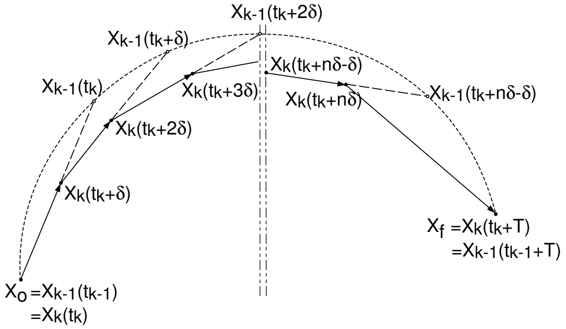

Fig. 1. Trajectories of a leader-follower pair during SLP. The dotted line represents the trajectory of the leader. Solid lines represent the trajectory of the follower. The dashed lines are trajectory segments which the follower plans but decides to alter, because of later measurements of the leader's state.

At time $t_k + \delta$ the follower reaches the state $x_k(t_k+\delta)$. Recalling that the trajectory driven by $u^*_{t_k}(\tau)$ is optimal from $x_k(t_k)$ to $x_{k-1}(t_k)$ and from Fact 3, we can divide the cost into two parts, one is actual and the other is planned, which are are both optimal with respect to their corresponding end points. That is

$$\begin{aligned} & \eta(x_k(t_k), x_{k-1}(t_k), \Delta, t_k, \Delta) = \\ = \;\; & \eta(x_k(t_k), x_{k-1}(t_k), \Delta, t_k, \delta) + \\ & + \eta(x_k(t_k+\delta), x_{k-1}(t_k), \Delta-\delta, t_k+\delta, \Delta-\delta) \end{aligned}$$

At time $t_k+\delta$, the follower updates its trajectory to catch up the leader at its new location $x_k(t_k+\delta)$. For this trajectory is optimal from $x_k(t_k+\delta)$ to $x_{k-1}(t_k+\delta)$ over time Δ, any path $x_k(t)$ $(t \in [t_k+\delta, t_k+\delta+\Delta])$ that is from $x_k(t_k+\delta)$ to $x_{k-1}(t_k+\delta)$ over time Δ and passes through $x_{k-1}(t_k)$ at time $t_k+\Delta = t_k+\delta+\Delta-\delta$ has equal or more cost. From Fact 2 it follows that:

$$\begin{aligned} & \eta(x_k(t_k+\delta), x_{k-1}(t_k+\delta), \Delta, t_k+\delta, \Delta) \\ \leq \;\; & \eta(x_k(t_k+\delta), x_{k-1}(t_k), \Delta-\delta, t_k+\delta, \Delta-\delta) + \\ & + C(x_{k-1}, t_k, \delta) \end{aligned}$$

From the last equation and the principle of optimality, we obtain

$$\begin{aligned} & C(x_k, t_k, 2\delta) \\ \leq \;\; & C(x_{k-1}, t_{k-1}, \Delta+\delta) - C(\hat{x}_k, t_k+2\delta, \Delta-\delta) \end{aligned}$$

We repeat this procedure until $t = t_k + n\delta$ where $\Delta + (n-1)\delta < T$ and $\Delta + n\delta \geq T$. Then

$$\begin{aligned} & C(x_k, t_k, n\delta) = \\ = \;\; & \sum_{i=0}^{n-1} \eta(x_k(t_k+i\delta), x_{k-1}(t_k+i\delta), \Delta, t_k+i\delta, \delta) \\ \leq \;\; & C(x_{k-1}, t_{k-1}, \Delta+(n-1)\delta) - \\ & - C(\hat{x}_k, t_k+n\delta, \Delta-\delta) \end{aligned} \tag{6}$$

When $t \in [t_k+n\delta, t_k+T]$, the leader reaches the final state and stays static. During this time period, no matter how many times the follower updates its movement, it will move on the same path that was determined at time $t = t_k + n\delta$. This path, which is indicated by the last solid line in Fig. 1, is locally optimal between the states $x_k(t_k+n\delta)$ and $x_k(t_k+T)$ over $T - n\delta$ units of time. Therefore

$$\begin{aligned} C(x_k, t_k+n\delta, T-n\delta) \leq C(\hat{x}_k, t_k+n\delta, \Delta-\delta) + \\ + C(x_{k-1}, t_k+(n-1)\delta, T-(n-1)\delta-\Delta) \end{aligned} \tag{7}$$

From (6), (7) we obtain

$$C(x_k, t_k, T) \quad \leq \quad C(x_{k-1}, t_{k-1}, T) \tag{8}$$

Writing $C_k = C(x_k, t_k, T)$ for convenience, we can see that $C_k \leq C_{k-1}$. Thus, C_k is bounded below and we conclude that $\lim_{k\to\infty} C_k$ exists. ■

Of course, the convergence of trajectories' cost does not imply the convergence of the trajectories themselves. If there exist multiple locally optimal trajectories connecting the leader and follower at any updating times, then the convergence of trajectories is not guaranteed. However, if we restrict the pursuit process to take place within a "small" region by selecting Δ sufficiently small, there will generally exist a unique locally optimal trajectory from the follower to the leader at every updating time $t_k + i\delta$, and the agents' trajectories converge:

Lemma 2: (Uniqueness of limiting trajectory): If at each updating time, the locally optimal trajectory obtained through SLP is unique, then the limiting trajectory $x_\infty(t)$ is also unique.

Proof: Suppose there exist more than one limiting trajectories, for example $x_1(t)$ and $x_2(t)$. Let $x_1(t) \neq x_2(t)$ for $t \in [t_1, t_2] \cup \ldots \cup [t_{n-1}, t_n]$. From Lemma 1, the two trajectories must have equal costs.

Let a leader $x_{k-1}(t)$ evolve along $x_1(t)$, while the follower $x_k(t)$ does so along $x_2(t)$. If no update occurs during $[t_1, t_2]$, then $x_2(t)$ costs less during $[t_1, t_2]$ because the follower moves along $x_2(t)$ and we have assumed that the optimal trajectories from follower to leader are unique. A similar argument on other intervals where $x_1 \neq x_2$ leads to the fact that the cost along $x_2(t)$ is less than that along $x_1(t)$ if no update occurs during $t \in [t_1, t_2] \cup [t_3, t_4] \ldots$. This contradicts the assumption that x_1 and x_2 have equal costs.

Next, suppose that the follower updates its trajectory once during $[t_1, t_2]$, as Fig. 2 illustrates. Separate the curves during $[t_1, t_2]$ into several segments (which have been labeled 1 through 5), and indicate the cost along curve i as C_i. From the uniqueness of local optimum, we have $C_1 + C_5 < C_3$ and $C_2 < C_5 + C_4$. Hence $C_1 + C_2 < C_3 + C_4$, which means $x_2(t)$ has less cost than $x_1(t)$ during $[t_1, t_2]$.

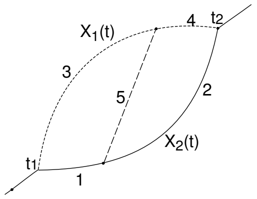

Fig. 2. Illustrating the case of a single trajectory update occurring at a point where the leader and follower trajectories differ.

A similar argument shows that if there are multiple updates during $[t_1, t_2]$, the cost along $x_2(t)$ is still less than that of $x_1(t)$. Iterating on the time intervals during which $x_1 \neq x_2$ leads to the conclusion that $x_2(t)$ costs less than $x_1(t)$, which is a contradiction. ■

The following definitions will be necessary for discussing the properties of the limiting trajectory.

*Definition 1: Let $\gamma_1(t)$ and $\gamma_2(t)$ be trajectories of (1), defined on time intervals I_1 and I_2 respectively, where $I_1 \cap I_2 \neq \phi$. We say that γ_1 and γ_2 **overlap** if $\gamma_1(t) = \gamma_2(t)$ for all $t \in I_1 \cap I_2$.*

*Definition 2: Let $\gamma_1(t)$ and $\gamma_2(t)$ be trajectories of (1), defined on a time interval I_1 and another time interval I_2 respectively, where $I_1 \cap I_2 \neq \phi$. The **composition** of $\gamma_1(t)$ and $\gamma_2(t)$ on the interval $I_1 \cup I_2$ is defined as*

$$\gamma_1 \circ \gamma_2 \triangleq \begin{cases} \gamma_1(t) & t \in I_1, t \notin I_2 - I_1 \cap I_2 \\ \gamma_2(t) & t \notin I_1, t \in I_2 - I_1 \cap I_2 \end{cases}$$

The locally optimal trajectories obtained at every updating time are smooth for many optimal control problems (e.g. the solution to the Euler-Lagrange equations). Nonetheless, $x_k(t)$ is only known to be piecewise smooth. However, we can show that the limiting trajectory is smooth in the entire interval $[0, T]$ if the locally optimal trajectories obtained at every updating time are smooth.

Lemma 3: Suppose that in Lemma 1 the updating interval δ and the following interval Δ satisfy that $0 < \delta < \Delta$, then for leader-follower pairs that evolve along the limiting trajectory, the planned trajectories $\hat{x}(t)$ and realized trajectories $x(t)$ overlap. Furthermore, if the locally optimal trajectories obtained at every updating time $t_k + i\delta$ are smooth, then the limiting trajectory is also smooth.

Proof: Consider an agent x_{k-1} that moves along the limiting trajectory $x_\infty(t)$. This implies that $x_{k-1}(t) = x_k(t + \Delta)$ for $\forall t \in [t_k, t_k + T]$. First we claim that in the time interval $[t_k + \delta, t_k + \Delta]$, the planned trajectory agrees with the realized one, i.e. $\hat{x}_k(t) = x_k(t), t \in [t_k + \delta, t_k + \Delta]$. Suppose that $\hat{x}_k(t) \neq x_k(t)$ for some $t \in [t_k + \delta, t_k + \Delta]$. Because $\hat{x}(t)$ is optimal from $x_k(t_k + \delta)$ to $x_k(t_k + \delta + \Delta)$, the trajectory

$$\bar{x}(t) = \begin{cases} \hat{x}_k(t) & t \in [t_k + \delta, t_k + \Delta) \\ x_k(t) & t \in [t_k + \Delta, t_k + \delta + \Delta] \end{cases}$$

has less cost than the trajectory $x_k(t)$ $(t \in [t_k + \delta, t_k + \delta + \Delta])$, which is updated by the follower at the time $t = t_k + \delta$ and is supposed to be optimal from $x_k(t_k + \delta)$ to $x_k(t_k + \delta + \Delta)$. Thus there is a contradiction. Hence we obtain $\hat{x}_k(t) = x_k(t)$ for $\forall t \in [t_k + \delta, t_k + \Delta]$. The same argument can be applied to other time periods.

Now, $\bar{x}(t)$ is smooth for $t \in [t_k, t_k + \Delta]$ because the local optima of (2) are smooth, and $x_k(t)$ is smooth for $t \in [t_k + \delta, t_k + \delta + \Delta]$ (second update step) for the same reason. Furthermore, we know that $\hat{x}_k(t) = x_k(t)$ for $\forall t \in [t_k + \delta, t_k + \Delta]$. Thus the actual trajectory $x_k(t)$ $(t \in [t_k, t_k + 2\delta])$ is smooth. Repeating this argument for $t \in [t_k + 2\delta, t_k + 3\delta]$, etc, leads to the result that the entire trajectory $x_k(t)$ $(t \in [t_k, t_k + T])$ is smooth. ■

Before proceeding to the main theorem, we will require that the optimal cost in (2) changes "little" with small changes to the endpoints of a trajectory:

Condition 1: Assume there exists an $\varepsilon > 0$ such that for all $a, b_1, b_2 \in \mathbb{D}$ and all $\Delta > 0$, the optimal cost $\eta(a, b_1, \Delta, 0, \Delta)$ from a to b_1 and $\eta(a, b_2, \Delta, 0, \Delta)$ from a to b_2 satisfy

$$\begin{aligned} & \|b_1 - b_2\|_\infty < \varepsilon \\ \Rightarrow \quad & \|\eta(a, b_1, \Delta, 0, \Delta) - \eta(a, b_2, \Delta, 0, \Delta)\|_\infty < \mathcal{L}\Delta \end{aligned}$$

for some constant $\mathcal{L}$ independent of Δ.

Piecewise-optimal trajectories are not necessarily optimal. However, the composition of overlapping optimal trajectories is locally optimal, if Condition 1 is satisfied.

Lemma 4: (Composition of Optimal Trajectories): Let $\gamma_1(t)$ and $\gamma_2(t)$ be overlapping locally optimal trajectories *defined on the intervals I_1 and I_2 respectively, where $I_1 \cap I_2 \neq \phi$. If Condition 1 is satisfied, then the composition $\gamma_1 \circ \gamma_2$ is locally optimal on $I_1 \cup I_2$.*

Proof: It is enough to show that if $x^*(t)$ $(t \in [0, t_1+\Delta_1])$ and $x^*(t)$ $(t \in [t_1, T])$ are two locally optimal trajectories, where $0 < t_1 < t_1 + \Delta_1 < T$, and Condition 1 is satisfied, then the trajectory $x^*(t), t \in [0, T]$ is a locally optimal.

Take $0 < \Delta \leq \Delta_1$. From principle of optimality, we have that $x^*(t)$ $(t \in [0, t_1 + \Delta])$ and $x^*(t)$ $(t \in [t_1, T])$ are two locally optimal trajectories with respect to their corresponding end points. If $x^*(t)$ $(t \in [0, T])$ is not a local optimum, there must exist an $\epsilon < \varepsilon$ and an optimum $x(t) \in \mathbb{D} \times [0, T]$ satisfying $\|x(t) - x^*(t)\|_\infty < \epsilon$ and $C(x(t), 0, T) < C(x^*(t), 0, T)$, as Fig. 3 illustrates.

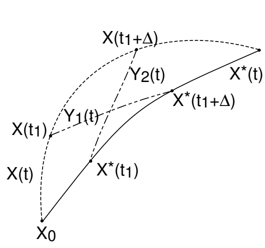

Fig. 3. Illustrating the proof of Lemma 4: overlapping optimal trajectories form a locally optimal trajectory.

Construct two optimal trajectories $y_1(t), y_2(t)$ $(t \in [t_1, t_1 + \Delta])$ connecting $x(t)$ and $x^*(t)$ such that $x^*(t_1) = y_2(t_1)$, $x^*(t_1 + \Delta) = y_1(t_1 + \Delta)$, $x(t_1) = y_1(t_1), x(t_1 + \Delta) = y_2(t_1 + \Delta)$. From the principle of optimality, $x^*(t)$ and $x(t)$ $(t \in [t_1, t_1 + \Delta])$ are both optimal with respect to their corresponding end points. Now from Condition 1, we have

$$\begin{aligned} C(y_1(t), t_1, \Delta) &< C(x(t), t_1, \Delta) + \mathcal{L}\Delta \\ C(y_2(t), t_1, \Delta) &< C(x^*(t), t_1, \Delta) + \mathcal{L}\Delta \end{aligned} \tag{9}$$

For $x^*(t)$ $(t \in [0, t_1 + \Delta])$ and $x^*(t)$ $(t \in [t_1, T])$ are two unique local optimal trajectories, we have

$$\begin{aligned} & C(x^*(t), 0, t_1) + C(x^*(t), t_1, \Delta) \\ < \quad & C(x(t), 0, t_1) + C(y_1(t), t_1, \Delta) \end{aligned} \tag{10}$$

$$\begin{aligned} & C(x^*(t), t_1, \Delta) + C(x^*(t), t_1 + \Delta, T - t_1 - \Delta) \\ < \quad & C(x(t), t_1 + \Delta, T - t_1 - \Delta) + C(y_2(t), t_1, \Delta) \end{aligned} \tag{11}$$

Combining (9)-(11) leads to

$$C(x^*(t), 0, T) < C(x(t), 0, T) + 2\mathcal{L}\Delta \tag{12}$$

The cost $C(x(t), 0, T)$ was assumed to be less than $C(x^*(t), 0, T)$, but if we choose Δ so that

$$0 < \Delta < \frac{C(x^*(t), 0, T) - C(x(t), 0, T)}{2\mathcal{L}}$$

we see that (12) cannot hold. This is a contradiction because Δ can be arbitrarily small. Hence $x^*(t)$ $(t \in [0, T])$ must be a local optimum. ■

The next theorem is an immediate consequence of the above lemmas.

Theorem 1: Suppose a group of agents $\{x_k\}$ evolve under sampled local pursuit and at each updating time $t = t_k + i\delta$, the locally optimal trajectory from $x_k(t)$ to $x_{k-1}(t)$ is unique. If the updating interval δ and following interval Δ satisfy $0 < \delta < \Delta$ and Condition 1 holds, then the trajectory sequence $\{x_k\}$ converges to a unique local optimum. Furthermore, if the locally optimal trajectories from each follower to its leader are smooth, the limiting trajectory is also smooth.

Proof: From Lemma 2, the limiting trajectory is unique. We know that $x_\infty(t)$ $(t \in [0, \Delta))$ and $x_\infty(t)$ $(t \in [\delta, \delta + \Delta))$ are locally optimal for the realized trajectory and planned trajectories overlap (Lemma 3). The optimality of $x_\infty(t)$ $(t \in [0, \delta + \Delta))$ follows from Lemma 4. Repeating this argument on $[i\delta, i\delta + \Delta]$ $(i = 0, 1, 2 \ldots)$ leads to the result that $x_\infty(t)$ $(t \in [0, T])$ is locally optimal. The proof of smoothness follows from a similar argument. ■

Remarks: SLP is a cooperative, decentralized algorithm for learning optimal controls/trajectories, starting from a feasible solution. Each agent is only required to calculate optimal trajectories from its own state to that of its leader. Because agents are separated by Δ time units as they leave x_0, each agent relies on local information only in order to follow its predecessor and requires no knowledge of the global geometry. There is no need for agents to exchange or "fuse" local maps. Agents do not need to communicate their choice of coordinate systems as they evolve, nor do they need to know the coordinates of x_f. While it is possible that a group of agents could disperse and construct a global map from local information, such an approach would require significantly more computation and communication than SLP. SLP solves the optimal control problem in "short pieces' which makes it appropriate for systems with short-range sensors (for example, in the case of a swarm of robots exploring unknown terrain). Each agent solves a finite number of instances of the optimal control problem, with initial and final states which are "close" to one another; if a closed form solution is not available, SLP generally requires significantly fewer computations compared to solving the problem (numerically) from x_0 to x_f. The above implies that SLP is most useful when the optimal control problem is easier to solve over "short" distances.

We have thus far assumed a countable infinity of agents. It is however possible to achieve similar results with a finite number of agents that apply SLP to reach x_f from x_0, then return to x_0 (perhaps using SLP again). The required modifications are straightforward but will not be discussed here because of space considerations. paper. Finally, SLP is not guaranteed to converge to the global optimum. The choice of agent separation and updating interval can affect whether the limiting trajectory is a local or a global optimum. Some interesting cases involving spaces with holes or obstacles are discussed in [16], [15].

IV. EXAMPLES

First, consider a group governed by $\ddot{x}(t) + x(t) = u(t)$ where we want to minimize $\int_0^1 \dot{x}^2(t)+u^2(t)dt$ with $x(0) = 0, x(1) = 1$ and $\dot{x}(0) = 0, \dot{x}(1) = 0$. Using $\Delta = 0.5,\ \delta = 0.25$, the iterated trajectories produced by SLP converged to the optimum, as illustrated in Fig. 4.

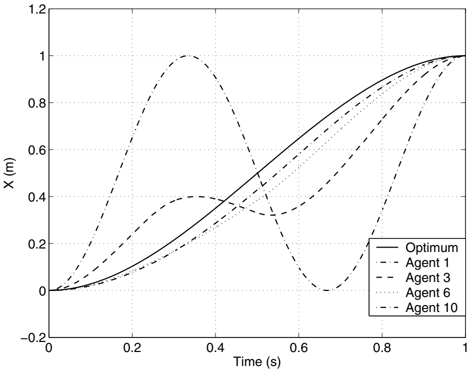

Fig. 4. SLP – first example. The initial trajectory was selected to be clearly sub-optimal; the iterated trajectories converge rapidly to the optimum using $\Delta = 0.5, \delta = 0.25$.

Next, consider a "geodesic discovery" problem in an environment consisting of a plane with two right cones, whose top view is shown in Fig. 5. The radii and heights of the cone were 800 and 1000 units of length, respectively. The agents were governed by $\dot{x}_k = u_k$ and were required to travel from $x_0 = (3500, 0, 0)$ to $x_f = (-1300, 0, 0)$. Minimum-length paths are difficult to compute in this setting because they involve optimal switching between different coordinate patches (those of the plane and the two cones). By applying SLP with $T = 1000,\ \Delta = 200,\ \delta = 1$, followers need to calculate locally optimal trajectories on at most two coordinate systems, thus reducing the complexity of the problem. Figure 5 illustrates the iterated trajectories.

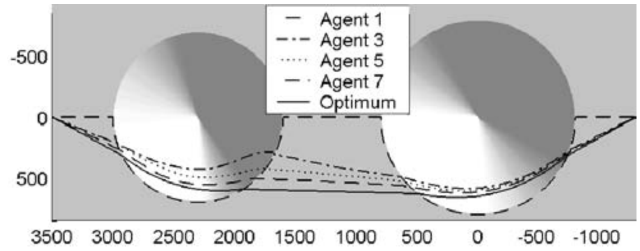

Fig. 5. SLP in a complex environment. The initial trajectory (along the borders of the cones) is easy to construct but far from optimal. The iterated trajectories (with $\Delta = 0.2T$) converge to the local optimum.

V. CONCLUSIONS AND ONGOING WORK

We explored a bio-inspired cooperative strategy, termed "Sampled Local Pursuit" (SLP), for solving a class of optimal control problems with fixed final time and state. SLP generalizes previous models that mimic the foraging behavior of ant colonies, and allows a collective to discover optimal controls, starting from an initial suboptimal solution. Members of the collective are only required to obtain local information on their environment and to calculate optimal trajectories to nearby neighbors. SLP relies on cooperation to perform a task which would be difficult or impossible for single systems, namely solving an optimal control problem with little information (in terms of coordinate systems that describe the environment or the coordinates of the final state) and short-range sensing. Extensions of this work could include broadening its scope to problems with free final time and state, and investigating SLP's convergence rate, as well as its ability to lead to global optima by choice of the algorithm's parameters.

REFERENCES

[1] S. Camazine, etc.(Eds), Self-Organization in Biological Systems, Princeton University Press, 2001.

[2] J. M. Pasteels and J-L. Deneubourg (Eds), From individual to collective behavior in social insects, Les Treilles Workshop, Basel, Boston, 1987.

[3] D. M. Gordon, Ants at work, The Free Press, New York, 1999.

[4] A. M. Bruckstein, Why the ant trails look so straight and nice, The Mathematical Intelligencer, 15(2), 1993, pp. 59-62.

[5] A. Jadbabaie, J. Lin and A. S. Morse, Coordination of groups of mobile autonomous agents using nearest neighbor rules, **AC**(48)-6, June 2003, pp. 988-1001.

[6] I. A. Wagner, M. Lindenbaum, A. M. Bruckstein, Distributed covering by ant-robots using evaporating traces, IEEE Trans. Robotics and Automation, Vol. 15, No. 5, 1999, pp. 918-933.

[7] K. Passino, M. Polycarpou, D. Jacques, M. Pachter, Y. Liu, Y. Yang, M. Flint and M. Baum, Cooperative control for autonomous air vehicles, in: R. Murphey, P. M. Pardalos(Eds), Cooperative control and optimization, Kluwer, 2002, pp. 233-272.

[8] S. I. Roumeliotis, G. A. Bekey, Distributed multi-robot localization, IEEE Tran. Robotics and Autom., Vol. 18, No. 5, 2002, pp. 781-795.

[9] R. Kurazume and S. Hirose, Study on cooperative positioning system: optimum moving strategies for CPS-III, Proc. 1998 IEEE Int'l Conf. Robotics and Automation, Vol. 4, May, 1998, pp. 2896-2903.

[10] M. Dorigo, V. Maniezzo, and A. Colorni, Ant systems: Optimization by a colony of cooperating agents, IEEE Trans. Systems, Man and Cybernetics, Part B, Vol 26, No. 1, 1996, pp. 29-41.

[11] A. M. Bruckstein, C. L. Mallows, and I. A. Wagner, Probabilistic pursuits on the grid, The American Mathematical Monthly, Vol. 104, No. 4, April, 1997, pp. 323-343.

[12] R. Fierro, P. Song, A. Das and V. Kumar, Cooperative control of robot formation, in: R. Murphey, P. M. Pardalos(Eds), Cooperative control and optimization, Kluwer, 2002, pp. 73-93.

[13] H. Yamaguchi, J. W. Burdick, Asymptotic stabilization of multiple nonholonomic mobile robots forming group formations, Proc. IEEE Int'l Conf. Robotics and Automation, Vol. 4, 1998, pp. 3573-3580.

[14] D. Hristu-Varsakelis, Robot formations: Learning minimum-length paths on uneven terrain, Proc. 8th IEEE Mediterranean Conf. Control and Automation, 2000.

[15] D. Hristu-Varsakelis and C. Shao, Biologically-inspired optimal control: Learning from social insects, International Journal of Control, vol.77, no. 18, p. 1549-66, Dec. 2004.

[16] C. Shao and D. Hristu-Varsakelis, Biologically-inspired algorithms for optimal control, Technical Report, Institute for Systems Research, University of Maryland, College Park, TR 2004-29, 2004.

[17] W. B. Dunbar and R. M. Murray, Model predictive control of coordinated multi-vehicle formations, Proc. 41st IEEE Conf. Decision and Control, p.4631-6, Dec. 2002.

[18] D. Q. Mayne, J. B. Rawlings, C. V. Rao and P. O. M. Scokaert, Constrained model predictive control: Stability and optimality, Automatica, vol. 36, p. 789-814, 2000.