52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

# A Generalized Reduced Gradient Method for the Optimal Control of Multiscale Dynamical Systems

Keith Rudd, Greg Foderaro, Silvia Ferrari

*Abstract*— **This paper considers the problem of computing optimal state and control trajectories for a multiscale dynamical system comprised of many interacting dynamical systems, or agents. A generalized reduced gradient (GRG) approach is presented for distributed optimal control (DOC) problems in which the agent dynamics are described by a small system of stochastic differential equations (SDEs). A new set of optimality conditions is derived using calculus of variations, and used to compute the optimal macroscopic state and microscopic control laws. An indirect GRG approach is used to solve the optimality conditions numerically for large systems of agents. By assuming a parametric control law obtained from the superposition of linear basis functions, the agent control laws can be determined via set-point regulation, such that the macroscopic behavior of the agents is optimized over time, based on multiple, interactive navigation objectives.**

## I. Introduction

This paper considers the problem of computing optimal state and control trajectories for a multiscale dynamical system comprised of many interacting dynamical systems, or agents. Many complex systems ranging from renewable resources [18] to very large scale robotic (VLRS) systems [17] can be described as multiscale dynamical systems comprised of many interactive agents. In recent years, significant progress has been made in formation control and stability analysis of teams of robots, or swarms, in which the mutual goal of the agents is to maintain a desired configuration, such as a triangle or a star formation, or a desired behavior, such as translating as a group (schooling), or maintaining the center of mass of the group stationary (flocking) [5], [7], [13], [17].While this literature has successfully illustrated that the behavior of large networks of interacting agents can be conveniently described and controlled by density functions, it has yet to provide an approach for controlling the agents such that their overall performance is optimized.

Recently, the authors proposed a coarse-grained optimal control approach for large, multiscale dynamical systems, referred to as distributed optimal control (DOC), that enables the optimization of density functions, and/or their moments, subject to the agents' dynamic constraints [6]. The DOC approach in [6] is applicable to multiscale dynamical systems comprised of many agents or processes that, on small spatial and time scales, are each described by a small set of ordinary differential equations (ODEs), referred to as the microscopic or *detailed* equations. On larger spatial and temporal scales, the agents' dynamics and interactions are assumed to give rise to macroscopic coherent behaviors, or *coarse dynamics*, described by partial differential equations (PDEs). This paper extends the capabilities of the DOC approach proposed in [6] for deterministic agent dynamics to agent dynamics that are governed by stochastic differential equations (SDEs). A new set of optimality conditions is derived, and a numerical indirect optimization approach is presented based on the GRG method [11].

In recent years, the optimal control of stochastic differential equations (SDEs) has gained increasing attention. Considerable research efforts have focused on the optimal control and estimation of SDEs driven by non-Gaussian processes, such as Brownian motion combined with Poisson processes, and various other stochastic processes [15], [21], [22]. The approach in [15], [21], [22] views the microscopic agent state as a random vector, and derives an SDE dynamic equation that involves the evolution of the statistics of the microscopic vector function, and may be integrated using stochastic integrals. Then, the performance of multiple agents can be expressed as an integral function of multiple, corresponding vector fields to be optimized subject to a set of SDEs. However, solutions can only be obtained for relatively few and highly idealized cases in which finite-dimensional, local approximations can be constructed, for example, via moment closure [21], [22]. Therefore, while optimal control of SDEs has been shown useful to selected applications in population biology and finance [15], [21], [22], it is yet to be successfully applied to multiscale systems in which the coarse dynamics do not obey these idealized conditions, and are instead dictated by realistic constraints (e.g., vehicle dynamics) and objectives (e.g., minimizing energy consumption, or maximizing coverage).

The GRG-DOC methodology presented in this paper relies on identifying a consistency relationship between the microscopic agent dynamics and a macroscopic description, such as the time-varying probability density function (PDF) of the agents' state. Unlike Nash Certainty Equivalence (NCE) or Mean Field methods, in which the (weak) couplings between agents are produced by the averaging of the microscopic agent dynamics and costs [4], [8], in the DOC approach the couplings need not be weak and may arise as a result of cooperative objectives expressed by the macroscopic cost function. Therefore, the cost function can represent objectives of a far more general form than NCE, and admit (optimal) solutions that entail strong couplings between the agent dynamics and control laws. Also, unlike prioritized and path-coordination methods [12], [25], the proposed DOC

K. Rudd, G. Foderaro, and S. Ferrari are with the Laboratory for Intelligence Systems and Controls (LISC), Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708-0005, {keith.rudd, greg.foderaro, sferrari}@duke.edu

978-1-4673-5716-6/13/$31.00 ©2013 IEEE

approach does not rely on decoupling the agents' dynamics, or on specifying the agents' distribution *a priori*. Instead, DOC optimizes the macroscopic behavior of the system subject to coupled microscopic agent dynamics, and relies on the existence of an accurate macroscopic evolution equation and an associated restriction operator that characterize the multiscale system to reduce the computational complexity of the optimal control problem. As a result, the computation required is far reduced compared to classical optimal control, and realizations of the trajectories of all agents over large spatial and time scales are calculated simultaneously without sacrificing optimality or completeness.

The paper is organized as follows. Section II describes the SDE-DOC problem formulation. The SDE-DOC optimality conditions and GRG-based numerical solution are presented in Section III. The effectiveness of the GRG-DOC approach is then demonstrated on a multi-agent trajectory optimization problem in Section IV.

## II. Problem Formulation

This paper considers the problem of computing optimal state and control trajectories for a multiscale dynamical system comprised of $N$ interacting dynamical systems, or agents. The dynamics of each agent on the microscopic scale can be described by a small system of the SDEs,

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{G}\mathbf{w}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{1}$$

where $\mathbf{x}(t) \in \mathcal{X} \subset \mathbb{R}^n$ is the microscopic state and $\mathbf{u} = c[\mathbf{x}(t), t] \in \mathcal{U} \subset \mathbb{R}^m$ is the microscopic control law, which is assumed to be a function of the state. The microscopic dynamics are influenced by additive Gaussian noise, where the disturbance, $\mathbf{w} \in \mathbb{R}^n$, is a vector of independent and identically distributed random variables from a standard Gaussian process, and $\mathbf{G}$ is a time-invariant matrix. A standard Gaussian process is used here for simplicity, but this approach is applicable to any diffusion process. It is assumed that the microscopic state, $\mathbf{x}$, of every agent is fully observable and error free.

On large spatial and temporal scales, the agents can be represented by a macroscopic state, denoted by $\mathbf{X} \in \mathbb{R}^\ell$, $\ell << n$, by means of a restriction operator. Depending on the macroscopic system performance to be optimized, the restriction operator may consist of the agent distribution and/or of its lower-order moments, such that $X(t) = \wp[\mathbf{x}(t), t]$ [10]. In this paper, the system restriction operator $\wp$ is assumed to be a time-varying probability density function (PDF), $\wp : \mathcal{X} \times \mathbb{R} \to \mathbb{R}$, such that the probability that the state of the $i^{th}$ agent has a value $\mathbf{x} \in B \subset \mathcal{X}$ is given by,

$$P[\mathbf{x}(t) \in B] = \int_B \wp[\mathbf{x}(t), t] d\mathbf{x} \tag{2}$$

Then, the agent PDF, $\wp$, is a non-negative probability function that must satisfy the normalization condition,

$$\int_{\mathcal{X}} \wp[\mathbf{x}(t), t] d\mathbf{x} = 1 \tag{3}$$

In many complex systems, such as autonomous vehicles and sensor networks, the performance to be optimized can be defined as an integral cost function of the macroscopic state $X$ and the microscopic control $\mathbf{u}$,

$$J = \int_{\mathcal{X}} \phi\left[X(t_f), t_f\right] d\mathbf{x} + \int_t \int_{\mathcal{X}} \mathscr{L}[X(t), \mathbf{u}(t), t] d\mathbf{x} dt, \tag{4}$$

where $\mathscr{L}$ is the Lagrangian function of the DOC problem. The multi-agent trajectory optimization problem considered in this paper seeks to determine the optimal trajectories for the macroscopic state $X^*$ and microscopic control $\mathbf{u}^*$ that minimizes the cost function (4), subject to the dynamic constraint (1) and the equality constraint (3).

Assuming that the agents exist only in the state space $\mathcal{X}$,

$$\wp[\mathbf{x}(t) \notin \mathcal{X}, t] = 0, \quad \forall \in (t_0, t_f] \tag{5}$$

and that no agents are created or destroyed, the evolution of the agent PDF, can be shown to be governed by the so-called advection-diffusion equation. The advection-diffusion equation is a parabolic PDE that describes the motion of a conserved scalar quantity, such as a PDF, as it is advected by a known velocity field and undergoes a diffusion process [3]. Since the agent distribution, $\wp$, is advected by a known velocity field $\mathbf{v} = \dot{\mathbf{x}} \in \mathbb{R}^n$, given by the detailed equation (1), and diffused by the additive Gaussian noise, the time-rate of change of $\wp$ can be defined as the sum of the negative divergence of the advection vector $(\wp\mathbf{v})$ and the divergence of diffusion vector $(\mathbf{G}\mathbf{G}^T\nabla\wp)$ [14]. Then, from the advection-diffusion equation, the agent PDF is governed by,

$$\begin{aligned} \frac{\partial \wp}{\partial t} &= -\nabla \cdot \{\wp[\mathbf{x}(t), t]\mathbf{v}(t)\} + \nabla \cdot \{(\mathbf{G}\mathbf{G}^T)\nabla\wp[\mathbf{x}(t), t]\} \\ &= -\nabla \cdot \{\wp[\mathbf{x}(t), t]\mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t]\} + \nu\nabla^2\wp[\mathbf{x}(t), t] \end{aligned} \tag{6}$$

where the $\nabla$ denotes a row vector of partial derivatives with respect to the elements of $\mathbf{x}$, and the diffusion coefficient is $\nu = \mathbf{G}\mathbf{G}^T$.

This paper presents a GRG method for computing the optimal trajectories of the macroscopic state $X^*$ and the microscopic control $\mathbf{u}^*$ that optimize $J$ over the time interval $(t_0, t_f]$. The optimization of $J$ is subject to the macroscopic dynamics (6), the normalization condition (3), and the state space constraint (5). and the initial and boundary conditions. Additionally, since the agents are assumed to exist in $\mathcal{X}$ at all times, their initial PDF, $g_0$ is typically given. Therefore, the optimization of $J$ is also subject to the initial and boundary conditions,

$$\wp[\mathbf{x}(t_0), t_0] = g_0(\mathbf{x}) \tag{7}$$

$$\{\nabla\wp[\mathbf{x}(t), t]\} \cdot \hat{\mathbf{n}} = 0, \quad \forall \in (t_0, t_f] \tag{8}$$

that require all agents to remain in $\mathcal{X}$ at all times, where $\hat{\mathbf{n}}$ is a vector normal to $\partial\mathcal{X}$ of unit length. The following section presents an indirect solution method based on a GRG approach for solving PDE-constrained optimization problems.

## III. Methodology

An indirect GRG solution method is presented in this section for computing the optimal macroscopic state and microscopic control trajectories for the DOC problem in (1)-(6). By this approach, a Lagrange multiplier, $\lambda(\mathbf{x},t)$, is used to adjoin the dynamic and equality constraints, (5)-(8), (3), to the integral cost function (4), obtaining the augmented integral cost function,

$$\hat{J} = \int_{\mathcal{X}} \phi\left\{X(t_f),t_f\right\} d\mathbf{x} + \int_t \int_{\mathcal{X}} \Bigg\{ \mathscr{L}\left[X(t),\mathbf{u}(t)\right] + \lambda(\mathbf{x},t)\Big[\frac{\partial \wp(\mathbf{x},t)}{\partial t} + \nabla\cdot\left[\wp(\mathbf{x},t)\mathbf{f}(\mathbf{x},\mathbf{u},t)\right] - \nu\nabla^2\wp(\mathbf{x},t)\Big]\Bigg\} d\mathbf{x}dt. \tag{9}$$

The necessary conditions for optimality for this augmented cost function are derived in Section III-A using the calculus of variations.

To have a closed form representation of the control for all $\mathbf{x}$, every element of the control vector, $u_j$, is parameterized as the sum of $m$ linearly-independent basis functions $\phi_1(\cdot),\ldots,\phi_m(\cdot)$, such that

$$u_j = \sum_k \phi_k(\mathbf{x})\alpha_{j,k}(t), \quad \text{for } j=1,\ldots,m. \tag{10}$$

Then, the goal of the DOC problem is to obtain the parameters, $\alpha^*_{j,k}(t)$, that minimize the cost function (4), subject to the aforementioned constraints. As shown in Section III-B, since the macroscopic state, $X$, and the Lagrangian multiplier, $\lambda$, can be found explicitly as a function of $\mathbf{u}$, a generalized reduced gradient (GRG) method [24] can be used to determine the optimal parameters of the control law (10).

### A. Optimality Conditions

The optimality conditions for the optimal control problem presented in Section II are derived here using calculus of variations. Let $\boldsymbol{\xi} = [\mathbf{u}^T, \wp, \lambda]^T$ denote a vector of variables for the DOC problem, where function arguments are omitted hereon for brevity. The necessary condition for optimality is,

$$\nabla\hat{J}(\mathbf{u},\wp,\lambda) = \lim_{\epsilon\to 0}\frac{\hat{J}(\boldsymbol{\xi}+\epsilon\delta\boldsymbol{\xi}) - \hat{J}(\boldsymbol{\xi})}{\epsilon} = 0, \tag{11}$$

where $\nabla\hat{J}$ is the gradient of $\hat{J}$ with respect to the variables, $\mathbf{u}$, $\wp$, $\lambda$, and the vector $\epsilon\ \delta\boldsymbol{\xi} = \epsilon\ [\delta\mathbf{u}^T, \delta\wp, \delta\lambda]^T$ contains the variations of the DOC variables.

The variation in the PDF, $\wp \to \wp + \epsilon\wp$, results in the condition,

$$\lim_{\epsilon\to 0}\frac{\hat{J}(\mathbf{u},\wp+\epsilon\delta\wp,\lambda) - \hat{J}(\mathbf{u},\wp,\lambda)}{\epsilon} = \int_{\mathcal{X}} \frac{\partial\phi}{\partial\wp}\Big|_{t_f}\delta\wp d\mathbf{x} + \int_t\int_{\mathcal{X}} \frac{\partial\mathscr{L}}{\partial\wp}\delta\wp + \lambda\left[\frac{\partial\delta\wp}{\partial t} + \nabla\cdot(\delta\wp\ \mathbf{f}) - \nabla^2\delta\wp\right] d\mathbf{x}dt = 0. \tag{12}$$

which provides the weak formulation of the DOC optimality conditions. The fundamental theorem of variational calculus (FTVC) is used to arrive at the strong formulation of the DOC optimality conditions. From the FTVC, and integration by parts, the partial derivatives acting on the variations are

$$\int_t\int_{\mathcal{X}} \lambda\frac{\partial\delta\wp}{\partial t}d\mathbf{x}dt = \int_{\mathcal{X}} \lambda\delta\wp d\mathbf{x}\Big|_{t_0}^{t_f} - \int_t\int_{\mathcal{X}} \frac{\partial\lambda}{\partial t}\delta\wp d\mathbf{x}dt, \tag{13}$$

$$\int_t\int_{\mathcal{X}} \lambda\nabla\cdot(\delta\wp\ \mathbf{f})\, d\mathbf{x}dt = \int_t\int_{\partial\mathcal{X}} \lambda(\mathbf{f}\cdot\hat{\mathbf{n}})\delta\wp d\mathbf{x}dt - \int_t\int_{\mathcal{X}} \nabla\lambda\cdot\mathbf{f}\delta\wp d\mathbf{x}dt, \tag{14}$$

$$\int_t\int_{\mathcal{X}} \nu\lambda\nabla^2\delta\wp d\mathbf{x}dt = \int_t\int_{\partial\mathcal{X}} \nu\lambda(\nabla\delta\wp\cdot\hat{\mathbf{n}})\delta\wp d\mathbf{x}dt - \int_t\int_{\mathcal{X}} \nu\nabla\lambda\cdot\nabla\delta\wp d\mathbf{x}dt = \int_t\int_{\partial\mathcal{X}} \nu\lambda(\nabla\delta\wp\cdot\hat{\mathbf{n}})\delta\wp d\mathbf{x}dt - \int_t\int_{\partial\mathcal{X}} \nu\nabla\lambda\cdot\hat{\mathbf{n}}\delta\wp d\mathbf{x}dt + \int_t\int_{\mathcal{X}} \nu\nabla^2\lambda\delta\wp d\mathbf{x}dt. \tag{15}$$

Because an initial condition for $\wp$ is given at $t_0$, as shown in (7), the initial variation in the PDF is $\delta\wp\big|_{t_0} = 0$, and (13) simplifies to

$$\int_t\int_{\mathcal{X}} \lambda\frac{\partial\delta\wp}{\partial t}d\mathbf{x}dt = \int_{\mathcal{X}} \lambda\delta\wp d\mathbf{x}\Big|_{t_f} - \int_t\int_{\mathcal{X}} \frac{\partial\lambda}{\partial t}\delta\wp d\mathbf{x}dt. \tag{16}$$

The boundary condition (8) implies that (15) simplifies to

$$\int_t\int_{\mathcal{X}} \nu\lambda\nabla^2\delta\wp d\mathbf{x}dt = -\int_t\int_{\partial\mathcal{X}} \nu\nabla\lambda\cdot\hat{\mathbf{n}}\delta\wp d\mathbf{x}dt + \int_t\int_{\mathcal{X}} \nu\nabla^2\lambda\delta\wp d\mathbf{x}dt. \tag{17}$$

Then, by substituting the results in (14), (16), and (17) into (12), and grouping like terms, the variation in (12) can be written as

$$0 = \int_{\mathcal{X}}\left(\frac{\partial\phi}{\partial\wp} + \lambda\right)\delta\wp\Big|_{t_f} d\mathbf{x} + \int_t\int_{\partial\mathcal{X}} \left(\lambda(\mathbf{f}\cdot\hat{\mathbf{n}}) + \nu\nabla\lambda\cdot\hat{\mathbf{n}}\right)\delta\wp d\mathbf{x}dt + \int_t\int_X\left(\frac{\partial\mathscr{L}}{\partial\wp} - \frac{\partial\lambda}{\partial t} - \nabla\lambda\cdot\mathbf{f} - \nu\nabla^2\lambda\right)\delta\wp d\mathbf{x}dt. \tag{18}$$

By the FTVC, the variation in (18) can be written as the adjoint PDE:

$$\frac{\partial\lambda}{\partial t} = \frac{\partial\mathscr{L}}{\partial\wp} - \nabla\lambda\cdot\mathbf{f} - \nu\nabla^2\lambda \tag{19}$$
$$\text{SJT: } \lambda(\mathbf{x},t_f) = -\frac{\partial\phi}{\partial\wp}\Big|_{t_f} \quad \mathbf{x}\in\mathcal{X},$$
$$\lambda(\mathbf{f}\cdot\hat{\mathbf{n}}) + \nu(\nabla\lambda)\cdot\hat{\mathbf{n}} = 0 \quad \mathbf{x}\in\partial\mathcal{X}$$

The variation in the control law, $\mathbf{u} \to \mathbf{u} + \epsilon\delta\mathbf{u}$,

$$\lim_{\epsilon\to 0} \frac{\hat{J}(\mathbf{u}+\epsilon\delta\mathbf{u}) - \hat{J}(\mathbf{u})}{\epsilon} = \tag{20}$$
$$\int_t \int_{\mathcal{X}} \frac{\partial \mathscr{L}}{\partial \mathbf{u}} + \lambda \left[ \nabla \cdot \left( \wp \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \delta\mathbf{u} \right) \right] d\mathbf{x}dt =$$
$$\int_t \int_{\mathcal{X}} \frac{\partial \mathscr{L}}{\partial \mathbf{u}} - \nabla\lambda \cdot \left( \wp \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right) \delta\mathbf{u} d\mathbf{x}dt +$$
$$\int_t \int_{\partial\mathcal{X}} \lambda \left( \wp \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \cdot \hat{\mathbf{n}} \right) \delta\mathbf{u} d\mathbf{x}dt.$$

must equal zero for optimality, by the FTVC, i.e.:

$$0 = \frac{\partial \mathscr{L}}{\partial \mathbf{u}} - \nabla\lambda \cdot \left( \wp \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right). \tag{21}$$

Finally, the variation in the Lagrange multiplier, $\lambda \to \lambda + \epsilon\delta\lambda$, leads to the macroscopic state equation. Thus, the DOC optimality conditions are given by the set of PDEs:

$$\frac{\partial \wp}{\partial t} = -\nabla \cdot (\wp \mathbf{f}) + \nu \nabla^2 \wp \tag{22}$$
$$\text{SJT: } \wp(\mathbf{x}, t_0) = p(\mathbf{x}) \quad \mathbf{x} \in \mathcal{X},$$
$$\nabla\wp \cdot \hat{\mathbf{n}} = 0 \quad \mathbf{x} \in \partial\mathcal{X}$$
$$\frac{\partial \lambda}{\partial t} = \frac{\partial \mathscr{L}}{\partial \wp} - \nabla\lambda \cdot \mathbf{f} - \nu\nabla^2\lambda \tag{23}$$
$$\text{SJT: } \lambda(\mathbf{x}, t_f) = -\frac{\partial \phi}{\partial \wp}\Big|_{t_f} \quad \mathbf{x} \in \mathcal{X},$$
$$\lambda = 0 \quad \mathbf{x} \in \partial\mathcal{X}$$
$$0 = \frac{\partial \mathscr{L}}{\partial \mathbf{u}} - \nabla\lambda \cdot \left( \wp \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right). \tag{24}$$

The macroscopic state (22) and adjoint (23) equations are parabolic PDEs. The control equation (24) is an algebraic equation relating the optimal $\mathbf{u}$ to $\wp$ and $\lambda$. If (22)-(24) are satisfied, then the resulting $\wp$ and $\mathbf{u}$ are the optimal control and resulting agent distribution for the macroscopic control problem. To obtain the sufficient conditions for optimality, the second-order variations of $\hat{J}$ may be tested to verify that these values in fact are at an extremal that is a minimum of $J$, but in this paper, the solutions are considered to be optimal if any perturbations only increase the value of $J$. The following subsection presents an GRG method to solve the optimality conditions to determine optimal DOC trajectories.

### *B. Numerical Solution Via GRG*

The DOC optimality conditions (22)-(24) consist of a coupled set of parabolic PDEs. Because analytical solutions to these PDEs are not available, this paper presents a GRG approach for reducing the computation required by the numerical solution of the DOC optimality conditions. The approach exploits the causality of the macroscopic dynamic equation (6) to represent $\hat{J}$ solely as a function of $\mathbf{u}$. Then an extremum of the DOC problem (1)-(6) can be found by determining the parameters of the control laws (10) that satisfy the optimality conditions.

GRG methods improve iteratively upon the approximation of the optimal control law and of the macroscopic state and co-state (or Lagrangian), by holding the other fixed during each update. During every iteration of the GRG algorithm, the latest approximation of $\mathbf{u}^* = \mathbf{c}^*[\mathbf{x}(t), t]$, in parameterized form (10) is used to solve macroscopic state and adjoint PDEs, (22) and (23), to obtain an approximation for $\wp^*$ and $\lambda^*$. Subsequently, holding the approximations of $\wp^*$ and $\lambda^*$ fixed, the approximation for $\mathbf{u}^*$ is updated so as to minimize (4), and satisfy the third and final optimality condition. This process is repeated until the norm of the gradient is below a user-defined tolerance or any update to $\mathbf{u}^*$ causes an increase in $J$.

The GRG method falls under a larger class of optimization techniques referred to as *Nested Analysis and Design* (NAND). In NAND approaches, the gradient is obtained at each iteration of the optimization by eliminating the state and co-state variables by solving the PDEs using a numerical algorithm, and only the control is considered [1]. Alternatively, a *Simultaneous Analysis and Design* (SAND), or full space, optimization strategy could be used in which the optimization over the state, co-state, and control are preformed simultaneously. However, it has been shown that SAND methods are often very ill conditioned, where the individual PDEs in the NAND techniques are typically better conditioned [2].

**Algorithm 1** GRG Optimality Solver

```
initialize α_{j,k}(t)
while ||g|| >TOL do
    φ̃ ← solve macroscopic state PDE (u)
    λ̃ ← solve adjoint PDE (φ̃, u)
    for all ℓ do
        g_ℓ ← compute gradient (φ̃, λ̃, u)
    end for
    for all j, k do
        α_{j,k} ← update α_{j,k} (J, g)
    end for
end while
```

An analytical representation of the gradient of the cost function $J$, denoted by $\mathbf{g}$, with respect to the controls $\mathbf{u}$ can be found, thereby circumventing the need for finite difference to approximate the gradient, greatly reducing the computational requirements. The gradient of $J$ is calculated as follows. Let $\tilde{\wp}$ and $\tilde{\lambda}$ satisfy (22) and (23), respectively, for a given $\mathbf{u}$. Then the gradient is given by

$$\nabla_u J = \nabla_u \hat{J}\Big|_{\tilde{\wp},\tilde{\lambda}} = \int_{\mathcal{X}} \frac{\partial \phi}{\partial \wp} \nabla_u \wp \; \delta\mathbf{u}\Big|_{\tilde{\wp},t_f} d\mathbf{x} + \tag{25}$$
$$\int_t \int_{\mathcal{X}} \left\{ \frac{\partial \mathscr{L}}{\partial \mathbf{u}} \delta\mathbf{u} + \frac{\partial \mathscr{L}}{\partial \wp} \nabla_u \wp \; \delta\mathbf{u} + \nabla_u \lambda \; \delta\mathbf{u} \left[ \frac{\partial \wp}{\partial t} + \nabla \cdot (\wp \mathbf{f}) \right] + \right.$$
$$\lambda \left[ \frac{\partial}{\partial t} (\nabla_u \wp \; \delta\mathbf{u}) + \nabla \cdot \left( \nabla_u \wp \mathbf{f} \delta\mathbf{u} + \wp \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \delta\mathbf{u} \right) \right]$$
$$\left. - \nu \nabla^2 \nabla_u \wp \delta\mathbf{u} \right\}_{\tilde{\wp},\tilde{\lambda}} d\mathbf{x}dt$$

Performing integration by parts and recalling that $\tilde{\wp}$ and $\tilde{\lambda}$

were defined to satisfy (22) and (23), equation (25) becomes

$$\nabla_u J = \int_t \int_{\mathcal{X}} \left[ \frac{\partial \mathscr{L}}{\partial \mathbf{u}} - \nabla \tilde{\lambda} \cdot \left( \tilde{\wp} \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right) \right] \delta \mathbf{u} d\mathbf{x} dt. \quad (26)$$

Let the time be discretized into $Q$ equally spaced points, $t_q = t_0 + q\Delta t$, where $q = 0, ..., Q$, and $\Delta t = (t_f - t_0)/Q$. Then from (26) it follows that

$$\left. \frac{\partial J}{\partial \alpha_{j,k}} \right|_{t=t_q} \approx \Delta t \int_X \left[ \frac{\partial \mathscr{L}}{\partial u_j} - \nabla \tilde{\lambda} \cdot \left( \tilde{\wp} \frac{\partial \mathbf{f}}{\partial u_j} \right) \right]_{t=t_q} \phi_k d\mathbf{x}. \quad (27)$$

The previous equation gives the gradient of the cost function with respect to the parameters that determine the control $\mathbf{u}$. Using this expression of the gradient, $\mathbf{u}$ can be updated using one of many gradient-based optimization schemes, such as Sequential Quadratic Programming (SQP). The algorithm for solving the optimality conditions is then given in Algorithm 1. The next section demonstrates the use of Algorithm 1 to find the optimal control law in a multi-agent path planning problem.

## IV. Multi-agent Trajectory Optimization

The GRG method presented in the previous sections is demonstrated here on a multi-agent trajectory optimization problem, that obeys the problem formulation in Section II. Consider a system of $N$ cooperative agents with microscopic dynamics given by a single integrator model for a point robot that was modified from the model proposed in [26],

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} + \sigma \mathbf{I}_2 \begin{bmatrix} \eta_x \\ \eta_y \end{bmatrix} \quad (28)$$

where $\mathbf{q} = [x\ y]^T$ denotes the configuration vector of the $i^{th}$ agent, and $x$ and $y$ are the $xy$-coordinates. The microscopic control vector of the $i^{th}$ agent is $\mathbf{u} = [v_x\ v_y]^T$, where $v_x$ and $v_y$ are linear velocities in the $x$ and $y$ directions, respectively. The disturbance vector is $\mathbf{w} = [\eta_x\ \eta_y]^T$, where $\eta_x$ and $\eta_y$ are independent random variables with values given by standard Gaussian processes, $\sigma$ is a constant, and $\mathbf{I}_2$ is the identity matrix. The agents operate in a workspace $\mathcal{W} = [L,\ 0] \times [0,\ L] \subset \mathbb{R}^2$ over a time interval $(t_0, t_f]$. The system restriction operator is a time-varying PDF of the agent states, $\wp : \mathcal{X} \times \mathbb{R} \to \mathbb{R}$, where $\wp(\mathbf{q}, t)$ provides the probability that the $i^{th}$ agent has the configuration $\mathbf{q}$ at time $t$. Then $\wp$ describes the density of the agents in the state space $\mathcal{X} = \mathcal{W}$.

The agents have the goal of traveling to a time-invariant target distribution, $p(\mathbf{x})$, that is known *a priori*, while minimizing the energy consumed through control. The objective function to be minimized can be written in terms of $\wp$, and is given by the integral cost function

$$J(\mathbf{u}) = \int_{\mathcal{W}} \phi(X)\Big|_{t_f} d\mathbf{x} + \int_{\mathcal{W}} \int_{t_0}^{t_f} \mathscr{L}(\mathbf{u}) dt d\mathbf{x} = \int_{\mathcal{W}} w_\wp (g - \wp)^2 \Big|_{t_f} d\mathbf{x} + \int_{\mathcal{W}} \int_{t_0}^{t_f} e^{\frac{w_u}{2}(v_x^2 + v_y^2)} dt d\mathbf{x} \quad (29)$$

where $w_\wp$ and $w_u$ are user-defined constant weights.

The optimal agent PDF $\wp^*$ and control $\mathbf{u}^*$ can be computed as follows. The control, $\mathbf{u}$, is approximated by the Fourier sine series

$$\mathbf{u}(\mathbf{x}, t) = \sum_{n=1}^{a} \sum_{m=1}^{b} \sin[n\pi(x_1+1)/2] \sin[m\pi(x_2+1)/2] \alpha_{mnj}(t). \quad (30)$$

This form ensures that $\mathbf{u} = 0$ on the boundary, forcing $\mathbf{f} \cdot \hat{\mathbf{n}} = 0$, which simplifies the boundary condition in (23) to $\nabla \lambda \cdot \hat{\mathbf{n}} = 0$. With this parameterized representation of the control, the gradient equation (27) is given by

$$\frac{\partial J}{\partial \alpha_{qpj}} \approx \Delta t \int_{\mathcal{W}} \left[ w_u u_j e^{\frac{w_u}{2}(v_{xi}^2 + v_{yi}^2)} - \wp \frac{\partial \lambda}{\partial x_j} \right] \times \sin[p\pi(x+1)/2] \sin[q\pi(y+1)/2] d\mathbf{x}. \quad (31)$$

where $u_j$ and $x_j$ denote the $j^{th}$ component of $\mathbf{u}$ and $\mathbf{x}$, respectively.

The numerical scheme used to solve (22) and (23) consists of a modified Galerkin method. A Galerkin type method was chosen for its non-dissipative property [9], [19]. In this modified approach, the solution is approximated by the linear combination of a Fourier basis and Gaussian radial basis functions (RBF), which are used to enforce the boundary conditions at each point of the integration.

An initial guess of $\alpha_{qpj} = 0$ was used to define the control for the first iteration of the optimization (Algorithm 1). Then the state and adjoint problems, (22) and (23), were solved. The control parameters, $\alpha_{qpj}(t_n)$, were then updated using gradient descent.

The agents' feedback control laws can be obtained from a set-point regulation method that uses the optimal agent PDF, $\wp^*$, and open-loop control, $\mathbf{u}^*$, that are found by solving the optimality conditions (22)-(24), as desired set-points [23]. The closed-loop control of each agent is computed independently, such that the control value at time $t$ of the $i^{th}$ agent, $\mathbf{u}(t)$, is determined to minimize the deviation between the observed agent distribution, denoted as $\hat{\wp}(t)$, and the optimal distribution $\wp^*(t)$, and the deviation between $\mathbf{u}(t)$ and the optimal open-loop control $\mathbf{u}^*(t)$ [16],

$$\mathbf{u}^*(t) = \min_{\mathbf{u}(t)} \int_t^{t+\delta t} \frac{1}{2} \left\{ [\wp^*(\mathbf{x}, t) - \hat{\wp}(\mathbf{x}, t)]^2 + \|\mathbf{u}^*(t) - \mathbf{u}(t)\|^2 \right\} dt \quad (32)$$

where $\| \cdot \|$ is the Euclidean norm, and $\delta t$ is a user-defined time increment. The observed agent distribution, $\hat{\wp}$, is calculated from the states of all agents using kernel density estimation with a standard Gaussian kernel [20]. The optimal feedback control $\mathbf{u}^*$ is updated at each timestep by minimizing (32) using one of several available quadratic programming algorithms [16]. In this paper, $\delta t$ is chosen to be small for simplicity, such that $\delta t << t_f - t_0$.

### A. Numerical Simulations

The GRG method presented in Section III is illustrated here through a numerical example where the optimal agent trajectories are calculated for a system of $N = 250$ agents

with microscopic dynamics governed by the single integrator model (28) with $\sigma = 0.01$. The agents exist in a workspace $\mathcal{W} = [0,\ L] \times [0,\ L]$, $L = 16$ km, over a time interval $(t_0,\ t_f]$, where $t_0 = 0$ and $t_f = 16$ hr. The agents have a given initial distribution $g_0$ shown in Figure 1, and the initial microscopic states and sampled from $g_0$. The system's objective is to minimize the integral cost (29) with $w_\varphi = 100$, $w_u = 6$, by travelling to a known target agent distribution $p$, illustrated in Figure 2, while minimizing the energy consumed by control. The solution to the trajectory optimization problem is found using the GRG approach, and the optimal agent distribution and microscopic states are plotted at four instants in time in Figure 3. It is seen from the results that the optimal agent distribution $\wp^*$ reaches the target distribution $p$.

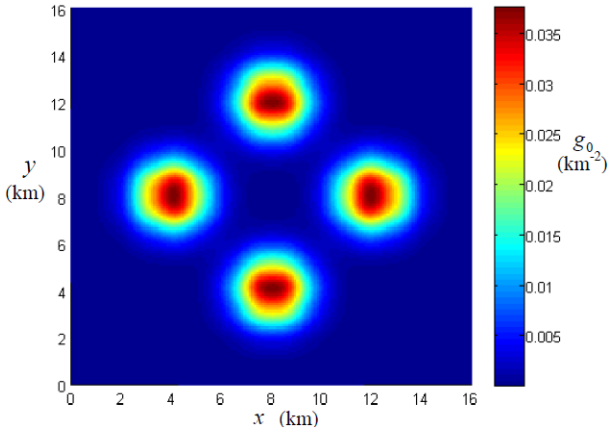


Fig. 1. Initial agent distribution, $g_0$.

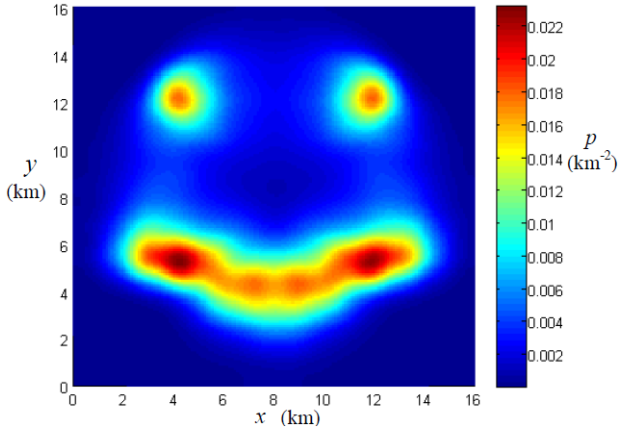


Fig. 2. Target agent distribution, $p$.

The agents' control input is given by the feedback control law (32) and calculated using MATLAB's quadratic program solver *quadprog*, where $\delta t = 20$ s. Then the microscopic states are updated by integrating the microscopic dynamic equations (28). The optimal microscopic state trajectories of $s = 50$ randomly-chosen agents are plotted in Figure 4, and the optimal microscopic control trajectories of $r = 3$ randomly-chosen agents are shown in Figure 5.

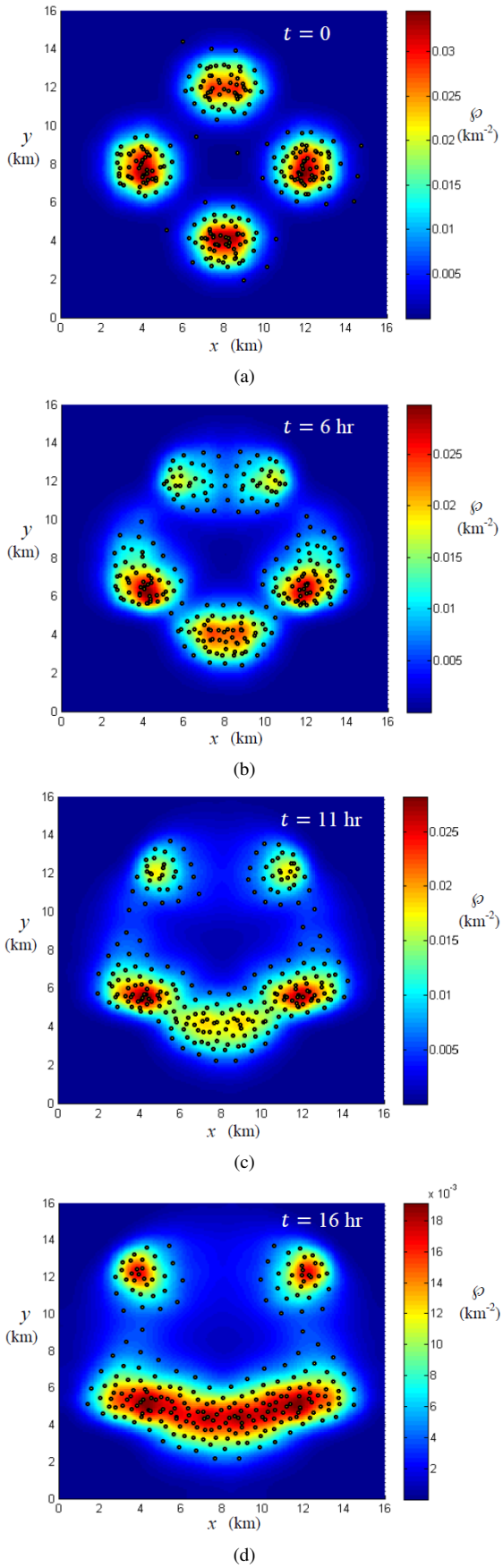


Fig. 3. Optimal evolution of agent distribution and microscopic states (yellow circles) for a system of $N = 250$ agents at four instants in time.

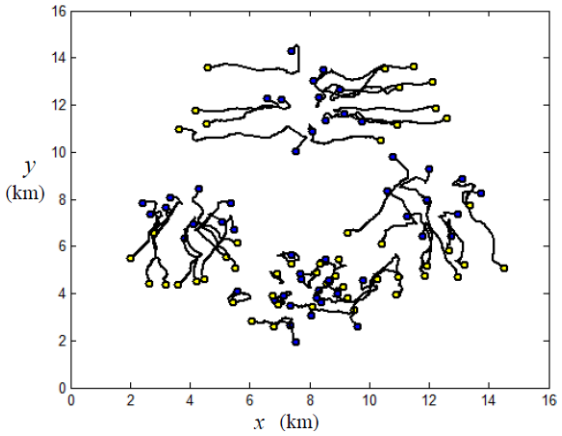


Fig. 4. Optimal microscopic state trajectories of $s = 50$ randomly-chosen agents traveling from their initial states (blue circles) to final states (yellow circles).

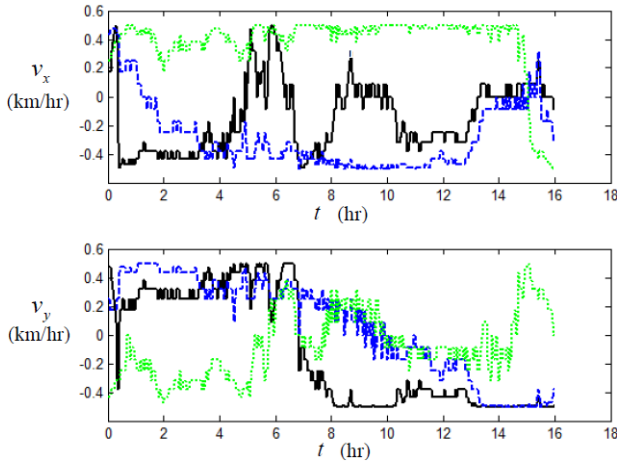


Fig. 5. Optimal microscopic control trajectories of $r = 3$ randomly-chosen agents.

## V. Conclusion

A GRG approach is presented to compute the optimal agent state and control trajectories for the DOC problem formulation with stochastic agent dynamics. This expands the capabilities of the DOC approach, which was previously only formulated for systems with deterministic agent dynamics. A new set of optimality conditions are derived for this case and are then solved using an indirect optimization method with GRG to obtain a functional representation of the optimal macroscopic state and microscopic open-loop control. A microscopic feedback control law is obtained using a set-point regulation method. The optimality conditions and the GRG approach are verified through a numerical simulation that determines the optimal state and control trajectories of a large system of agents with dynamics governed by a single integrator point robot model.

## References


[1] L.T. Biegler. *Large-Scale Pde-Constrained Optimization*. Lecture Notes in Computational Science and Engineering, 30. Springer Verlag, 2003.

[2] G. Biros and O. Ghattas. Parallel lagrange–newton–krylov–schur methods for pde-constrained optimization. part i: The krylov–schur solver. *SIAM Journal on Scientific Computing*, 27(2):687–713, 2005.

[3] J. P. Boyd. *Chebyshev and Fourier Spectral Methods, II Ed.* Dover, New York, NY, 2001.

[4] P.E. Caines. Mean field stochastic control. *Bode Lecture at the 48th IEEE Conf. on Decision and Control and 28th Chinese Control Conf.*, 2009.

[5] L. Pollini F. Giulietti and M. Innocenti. Autonomous formation flight. *IEEE Control Systems Magazine*, 20:3444–3457, 2000.

[6] G. Foderaro and S. Ferrari. Necessary conditions for optimality for a distributed optimal control problem. *Proc. IEEE Conference on Decision and Control, Atlanta, Georgia*, 2010.

[7] V. Gazi and K.M. Passino. Stability analysis of social foraging swarms. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 34(1):539 –557, feb. 2004.

[8] M. Huang, R.P. Malhame, and P.E. Caines. Large population stochastic dynamic games: Closed-loop mckean-vlasov systems and the nash certainty equivalence principle. *Communications in Information and Systems*, 6(3):221–252, 2006.

[9] M.R Kaazempur-Mofrad and C.R Ethier. An efficient characteristic galerkin scheme for the advection equation in 3-d. *Computer Methods in Applied Mechanics and Engineering*, 191(46):5345 – 5363, 2002.

[10] I.G. Kevrekidis, C.W. Gear, J.M. Hyman, P.G. Kevrekidis, O. Runborg, , and C. Theodoropoulos. Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis. *Communications in Mathematical Sciences*, 1(4):715–762, 2003.

[11] L.S. Lasdon, R.L. Fox, and M.W. Ratner. Nonlinear optimization using the generalized reduced gradient method. *Tech. Memo. No. 325, Department of Operations Research, Case Western Reserve University*, 1973.

[12] S.M. LaValle and S.A. Hutchinson. Optimal motion planning for multiple robots having independent goals. *Robotics and Automation, IEEE Transactions on*, 14(6):912 –925, dec 1998.

[13] N. E. Leonard and E. Fiorelli. Virtual leaders, artificial potentials and coordinated control of groups. In *Proceedings of the Conference on Decision and Control*, page 29682973, 2001.

[14] X. Mao. *Stochastic Differential Equations and Applications*. Horwood Publishing, 1997.

[15] S. Peng and Z. Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization*, 37(3):825843, 1999.

[16] S. J. Qin and T. A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

[17] J. H. Reif and H. Wang. Social potential fields: A distributed behavioral control for autonomous robots. *Robotics and Autonomous Systems*, 27(3):171–194, 1999.

[18] J. Sanchirico and J. Wilen. Optimal spatial management of renewable resources: Matching policy scope to ecosystem scale. *Journal of Environmental Economics and Management*, 50, 2005.

[19] T.K. Sengupta, S.B. Talla, and S.C. Pradhan. Galerkin finite element methods for wave problems. *Sadhana*, 30, 2005.

[20] J. S. Simonoff. *Smoothing Methods in Statistics*. Springer, 1996.

[21] A. Singh and J. P. Hespanha. Moment closure techniques for stochastic models in population biology. *IEEE Proceedings of the American Control Conference*, pages 4730–4735, 2006.

[22] A. Singh and J. P. Hespanha. A derivative matching approach to moment closure for the stochastic logistic model. *Bulletin of Mathematical Biology*, 69(6):1909–1925, 2007.

[23] R. F. Stengel. *Optimal Control and Estimation*. Dover Publications, Inc., 1986.

[24] W. Sun and Y.X. Yuan. *Optimization Theory and Methods: Nonlinear Programming*. Springer Optimization and Its Applications Series. Springer Science+Business Media, LLC, 2006.

[25] S. Thrun, M. Bennewitz, and W. Burgard. Finding and optimizing solvable priority schemes for decoupled path planning techniques for teams of mobile robots. *Robotics and Autonomous Systems*, 41(2):89–99, 2002.

[26] M.M. Zavlanos, M.B. Egerstedt, and G.J. Pappas. Graph-theoretic connectivity control of mobile robot networks. *Proceedings of the IEEE*, 99(9):1525 –1540, sept. 2011.