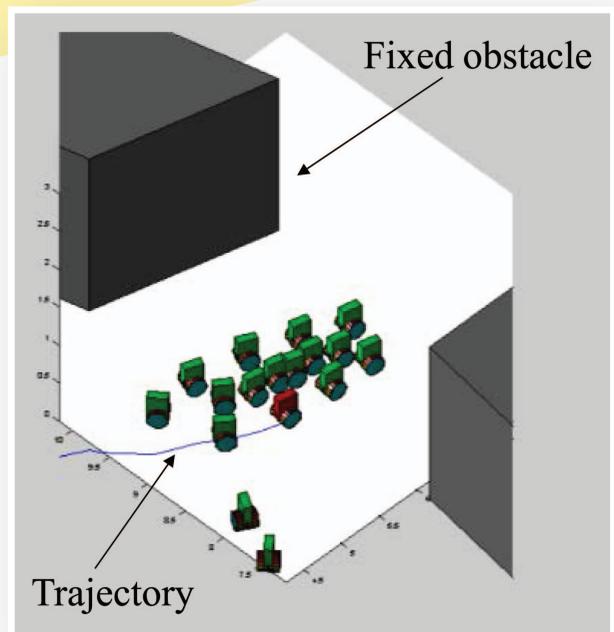


Distributed Optimal Control of Multiscale Dynamical Systems

A TUTORIAL

SILVIA FERRARI, GREG FODERARO,
PINGPING ZHU,
and THOMAS A. WETTERGREN



Many complex systems, ranging from renewable resources [1] to very-large-scale robotic systems (VLRs) [2], can be described as multiscale dynamical systems comprising many interactive agents. In recent years, significant progress has been made in the formation control and stability analysis of teams of agents, such as robots, or autonomous vehicles. In these systems, the mutual goals of the agents are, for example, to maintain a desired configuration, such as a triangle or a star formation, or to perform a desired behavior, such as translating as a group (schooling) or maintaining the center of mass of the group (flocking) [2]–[7]. 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 optimizing the agent density functions such that their mutual goals are optimized.

Digital Object Identifier 10.1109/MCS.2015.2512034
Date of publication: 17 March 2016

TABLE 1 Distributed optimal control notation.

Symbol	Meaning
i	Agent index
\mathbf{x}	Agent microscopic state
\mathbf{u}	Agent control
\mathcal{U}	Admissible control space
ρ	Agent distribution
ρ^*	Optimal agent distribution
$\hat{\rho}$	Estimated agent distribution
J	Integral cost function
\mathcal{L}	Lagrangian function
ϕ	Terminal cost function
T_0	Initial time
T_f	Final time
∇	Gradient (column) vector
f_i	j th Gaussian mixture component
ω_j	Weight of the j th Gaussian mixture component
μ_j	Mean of the j th Gaussian mixture component
\sum_j	Covariance matrix of the j th Gaussian mixture component
z	Number of Gaussian mixture components
ρ_k	FV approximation of advection operator
$\Delta t, \Delta \mathbf{x}$	Temporal and spatial discretization intervals
K, L	Number of temporal and spatial collocation points
λ	Lagrange multiplier or costate vector
K	Kernel function
$\mathbf{H}_{i,k}, \gamma_{i,k}$	Bandwidth matrix and weighting coefficient of kernel k stored by agent i
U	Potential function
U_{rep}	Repulsive potential function

This article describes a coarse-grained optimal control approach for multiscale dynamical systems, referred to as distributed optimal control (DOC), which enables the optimization of density functions, and/or their moments, subject to the agents' dynamic constraints. (See Table 1 for DOC notation.) The DOC approach is applicable to systems, such as swarms or teams, composed of many collaborative agents 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 agent behaviors and interactions are assumed to give rise to macroscopic coherent behaviors, or *coarse dynamics*, described by partial differential equations (PDEs) that are possibly stochastic.

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, or other stochastic processes [8]–[10]. In these efforts, the microscopic agent state is viewed as a random vector, and the dynamic equation takes the form of an SDE that describes the evolution of the statistics of the microscopic vector function and may be integrated using stochastic integrals. Then, the performance of N agents is expressed as an integral function of N vector fields to be optimized subject to N 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 [8], [9]. While the optimal control of SDEs has been primarily shown useful to selected applications in population biology and finance [8]–[10], recently it has also been successfully applied to obtain equilibrium strategies for multiagent systems (MAS) that obey a game-theoretic framework where each agent has its own cost function and the macroscopic system performance is expressed by a predefined collective potential function (see [11] and the references therein).

Other approaches that have been proposed for tackling the control of multiscale dynamical systems and provide practical yet tractable solutions even when the number of agents is large include prioritized planning techniques [12] and path-coordination methods [13], which first plan the agent trajectories independently and then adjust the microscopic control laws to avoid mutual collisions. Behavior-based control methods seek feasible solutions by programming a set of simple behaviors for each agent and by showing that the agents' interactions give rise to a macroscopic behavior, such as dispersion [2]. Swarm-intelligence methods, such as foraging and schooling [5], [14], [15], view each agent as an interchangeable unit subject to local objectives and constraints through which the swarm can converge to a range of predefined distributions or a satisfactory strategy. When the agents' dynamics and costs are weakly coupled, useful decentralized control strategies for multiscale dynamical systems can be obtained via the Nash certainty equivalence (NCE) or mean field principle [16], [17].

Similarly to DOC, NCE methods rely on identifying a consistency relationship between the microscopic agent dynamics and a macroscopic description, which in the NCE is the mass of the agents. However, while in NCE methods the (weak) couplings

between agents are produced by averaging the microscopic agent dynamics and costs, in DOC couplings need not be weak, and may arise as a result of collaborative objectives expressed by the macroscopic cost function. As a result, the cost function can represent objectives of a far more general form than other methods and admit solutions that entail strong couplings between the agent dynamics and control laws. Furthermore, while NCE methods rely on a macroscopic state description defined in terms of the expectation (or mean) of the evolving agent distribution, the DOC approach is applicable to other macroscopic descriptions, such as the agent probability density function (PDF) and its moments, thereby admitting a wide range of collaborative behaviors and objectives.

Unlike prioritized and path-coordination methods [12], [13], the DOC approach does not rely on decoupling the agent behaviors, or on specifying the agent 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 [18], and coupled agent objectives and control laws can be considered over large spatial and time scales without sacrificing optimality or completeness [19].

This tutorial provides an introductory overview of the DOC problem and of how it compares to existing optimal control methods. Two existing DOC solution methods are described, and the necessary and sufficient conditions for optimality are reviewed in the context of direct and indirect DOC methods. Similarly to classical optimal control, indirect DOC methods seek to determine solutions to the optimality conditions [20]. Direct DOC methods discretize the original problem formulation, in this case both with respect to space and time, to obtain a mathematical program that, in the general case, takes the form of a nonlinear program (NLP). Subsequently, each agent can compute its feedback control law based on the optimal time-varying agent PDF determined by DOC and on the actual agent distribution, which may be obtained via kernel density estimation. Finally, the applicability of the DOC approach is illustrated through a multiagent formation and path-planning application and an image-reconstruction problem.

BACKGROUND ON OPTIMAL CONTROL

Optimal control can be considered the most general approach to optimizing the performance of a dynamical system over time. Over the years, it has been successfully applied to control a wide range of processes, including physical, chemical, economic, and transportation systems. The classical optimal control formulation considers a system whose dynamics can be approximated by a small system of SDEs

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t), \mathbf{u}(t), t] + \mathbf{G}(t)\mathbf{w}(t), \quad \mathbf{x}(T_0) = \mathbf{x}_0, \quad (1)$$

where \mathbf{x} and \mathbf{u} are the system state and control input vectors, respectively, and $\mathbf{w} \in \mathbb{R}^s$ is a vector of random inputs with Gaussian distribution and zero mean [21]. The dynamics in (1) also depend on system parameters that represent the physical characteristics of the system and are expressed in units that scale both the inputs and outputs to comparable magnitudes. The term $\mathbf{G}(t)\mathbf{w}(t)$ is typically assumed to capture the random effects associated with parameter variations and uncontrollable inputs known as disturbances [21, p. 422].

Classical optimal control seeks to determine the state and control trajectories that optimize an integral cost function

$$J = \phi[\mathbf{x}(T_f)] + \int_{T_0}^{T_f} \mathcal{L}[\mathbf{x}(t), \mathbf{u}(t), t] dt, \quad (2)$$

over a time interval $(T_0, T_f]$, subject to (1) and, potentially, to an d -dimensional inequality constraint

$$\mathbf{q}[\mathbf{x}(t), \mathbf{u}(t), t] \leq \mathbf{0}_{d \times 1}. \quad (3)$$

When random effects can be neglected, optimal control problems in the form (1)–(3) can be approached by solving the necessary conditions for optimality known as Euler–Lagrange equations, which may be derived using calculus of variations [21]. Another approach is to use the recurrence relationship of dynamic programming to iteratively approximate candidate optimal trajectories, known as extremals, by optimizing the value function or cost-to-go. For a nonlinear system and a general cost function, the Euler–Lagrange equations amount to a Hamiltonian boundary-value problem for which there typically are no closed-form solutions. In this case, the optimal control problem is typically solved numerically using direct or indirect methods [21]–[23].

Indirect methods solve a nonlinear multipoint boundary value problem to determine candidate optimal trajectories. This requires deriving analytical expressions for the necessary conditions for optimality and then implementing a root-finding algorithm to solve the optimality conditions numerically. Direct methods determine near-optimal solutions by discretizing the continuous-time problem about collocation points and then transcribing it into a finite-dimensional NLP. A numerical optimization algorithm, such as sequential quadratic programming (SQP), is then used to find the optimal dynamic state and control variables directly from the discretized optimal control problem [24], [25]. Direct methods are typically easier to implement than indirect methods because they do not involve the derivation of analytical expressions, which may be challenging especially for nonlinear systems, have better convergence properties, and do not require initial guesses for the adjoint variables, which can be difficult to provide in the presence of inequality constraints [26]–[28].

When random effects are too important to be neglected, the stochastic principle of optimality can be obtained by

taking the expectation of the integral cost (2), ultimately reducing the stochastic optimal control problem to the minimization of a value function that differs from its deterministic counterpart by a term containing the spectral density matrix of the random input [21, pp. 422–432]. Traditionally, stochastic optimal control has dealt with a single dynamical system modeled by a small system of ODEs with additive Gaussian random inputs, such as (1). In recent years, however, considerable efforts have gone into the optimal control and estimation of stochastic differential equations driven by non-Gaussian processes such as Brownian motion and Poisson processes [8]–[11]. Let the state of the i th agent be represented by a random vector, \mathbf{x}_i , subject to a stochastic process that takes the form of the SDE

$$\begin{aligned} d\mathbf{x}_i(t) &= \{\mathbf{f}[\mathbf{x}_i(t), t] + \mathbf{B}\mathbf{u}_i(t)\}dt + \mathbf{G}[\mathbf{x}_i(t), t]d\mathbf{w}_i(t), \\ \mathbf{x}_i(t_0) &= \mathbf{x}_{i_0}, \end{aligned} \quad (4)$$

where \mathbf{w}_i denotes a Brownian motion process [11]. Then, the performance of N agents can be expressed as an integral function of $\mathbf{f}[\mathbf{x}_1(t), \dots, \mathbf{f}[\mathbf{x}_N(t)]$, and must be optimized subject to N SDEs in the form of (4), which can be integrated using stochastic integrals. While this approach allows modeling and optimizing stochastic processes that are not well described by ODEs with additive Gaussian distributions, it does not resolve the computational challenges associated with many collaborative agents. In fact, the optimal control of SDEs, such as (4), typically requires solving optimality conditions that are numerically more challenging than the Euler–Lagrange equations and do not afford any computational savings.

In many MAS, the goal is to control processes that can each be described by the ODEs in (1) but, because of common goals and objectives, ultimately lead to coupled optimality conditions and emergent behaviors governed by PDEs. To date, the control of PDEs or distributed-parameter systems has focused primarily on parabolic models, such as

$$\begin{aligned} \frac{\partial X}{\partial t} &= \mathcal{N}\{X(\mathbf{z}, t)\} + U(\mathbf{z}, t), \quad \mathbf{z} \in D, t \geq 0, \\ \mathcal{B}\{X(\mathbf{z}, t)\} &= 0, \quad \mathbf{z} \in \partial D, X(\mathbf{z}, 0) = X_0(\mathbf{z}), \quad \mathbf{z} \in D, \end{aligned} \quad (5)$$

where $\mathbf{z} = [x \ y]^T$ are the coordinates of the spatial domain $D \in \mathbb{R}^2$ with boundary denoted by ∂D , $\mathcal{N}\{\cdot\}$ and $\mathcal{B}\{\cdot\}$ are well-posed linear or nonlinear spatial differential operators, X_0 is a known function, and $U(\mathbf{z}, t)$ is the forcing function or control input to be determined. In a typical distributed-parameter system, the scalar input U can be used to control the spatial and temporal evolution of the state X and can be designed to be a function of X via feedback [29]. Thus, optimal control of PDEs seeks to optimize an integral cost function of X and U over a domain D , subject to a PDE such as (5). The solution to this class of distributed control problems can be obtained from the Karush–Kuhn–Tucker (KKT) conditions, as described in detail in [29].

The methods reviewed above show that optimal control is a promising framework for controlling the behavior of

MAS because it enables the optimization of integral cost functions in dynamic and uncertain settings. Like many of the aforementioned methods, DOC seeks to extend optimal control theory to new problem formulations that are not restricted to systems described by ODEs. In particular, DOC addresses the optimization of restriction operators that capture the macroscopic performance of a multiscale dynamical system by means of a reduced-order representation of the agents' state that is governed by PDEs. The DOC methods reviewed in this article focus on the optimal control of time-varying PDFs subject to a parabolic PDE known as the advection-diffusion equation, which can be obtained from (5) by considering multivariable state and control and vector-valued operators $\mathcal{N}\{\cdot\}$ and $\mathcal{B}\{\cdot\}$. Unlike existing methods for the control of parabolic PDEs, DOC considers multiscale systems in which agents have microscopic control inputs and dynamics that influence the advection and diffusion terms in (5), rather than optimizing the macroscopic state and control, X and U . As a result, DOC can be used to obtain agent control laws that optimize the macroscopic performance subject to the microscopic agent dynamics and constraints, as shown in the following sections.

DISTRIBUTED OPTIMAL CONTROL PROBLEM

Consider the problem of controlling a group or team of N collaborative dynamical systems, referred to as agents. Assume every agent can be described by a small system of SDEs in the form

$$\dot{\mathbf{x}}_i(t) = \mathbf{f}[\mathbf{x}_i(t), \mathbf{u}_i(t), t] + \mathbf{G}\mathbf{w}_i(t), \quad \mathbf{x}_i(T_0) = \mathbf{x}_{i_0}, \quad (6)$$

where $\mathbf{x}_i \in \mathcal{X} \subset \mathbb{R}^n$ and $\mathbf{u}_i \in \mathcal{U} \subset \mathbb{R}^m$ denote the microscopic agent state and control, respectively, \mathcal{X} denotes the microscopic state space, \mathcal{U} is the space of admissible microscopic control inputs, and \mathbf{x}_{i_0} is the initial conditions. Random parameter variations and disturbance inputs are modeled by an additive Gaussian disturbance vector $\mathbf{w}_i \in \mathbb{R}^s$ and the constant matrix $\mathbf{G} \in \mathbb{R}^{s \times s}$.

It is assumed that the agents cooperate toward one or more common objectives by virtue of the microscopic control such that, at large spatial and temporal scales, their performance over a time interval (T_0, T_f) can be expressed as an integral cost function of \mathbf{u}_i and a macroscopic state variable $X(t) = p(\mathbf{x}_i, t)$

$$J = \phi[p(\mathbf{x}_i, T_f)] + \int_{T_0}^{T_f} \int_{\mathcal{X}} \mathcal{L}[p(\mathbf{x}_i, t), \mathbf{u}_i(t), t] d\mathbf{x}_i dt, \quad (7)$$

where $p: \mathcal{X} \times \mathbb{R} \rightarrow \mathbb{R}$ is a restriction operator [30]. Based on the literature on swarm intelligence, MAS, and sensor networks, in this article the restriction operator is chosen to be a time-varying PDF such that, at any time t , the probability that the agent state \mathbf{x}_i is in a subset $B \subset \mathcal{X}$ is

$$P(\mathbf{x}_i \in B) = \int_B p(\mathbf{x}_i, t) d\mathbf{x}_i, \quad (8)$$

where p is a nonnegative function that satisfies the normalization property

$$\int_{\mathcal{X}} p(\mathbf{x}_i, t) d\mathbf{x}_i = 1, \quad (9)$$

and $Np(\mathbf{x}_i, t)$ is the density of agents in \mathcal{X} .

Then, assuming $\mathbf{x}_i(t) \in \mathcal{X}$ for all t and all i , the macroscopic dynamics of the multiscale system can be derived from the continuity equation as follows. From the detailed equation (6), the agent PDF p can be viewed as a conserved quantity advected by a known velocity field $\mathbf{v}_i(t) = \mathbf{f}[\mathbf{x}_i(t), \mathbf{u}_i(t), t]$ and diffused by the additive Gaussian noise $\mathbf{G}\mathbf{w}_i$ [31]. From the continuity equation and Gauss's theorem, the time-rate of change of p can be defined as the sum of the negative divergence of the advection vector ($p\mathbf{v}_i$) and the divergence of diffusion vector ($\mathbf{G}\mathbf{G}^T \nabla p$) [32]. Then, the evolution of the agent PDF is governed by the advection-diffusion equation

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\nabla \cdot [p(\mathbf{x}_i, t)\mathbf{v}_i(t)] + \nabla \cdot [(\mathbf{G}\mathbf{G}^T)\nabla p(\mathbf{x}_i, t)] \\ &= -\nabla \cdot [p(\mathbf{x}_i, t)\mathbf{f}(\mathbf{x}_i, \mathbf{u}_i, t)] + \nu \nabla^2 p(\mathbf{x}_i, t), \end{aligned} \quad (10)$$

which is a parabolic PDE in the form of (5), where ν is the diffusion coefficient, the gradient ∇ denotes a row vector of partial derivatives with respect to the elements of \mathbf{x}_i , and (\cdot) denotes the dot product.

Because the agent initial conditions are typically given, the initial agent distribution is a known PDF, $p_0(\mathbf{x}_i)$, and the macroscopic evolution equation (10) is subject to the initial and boundary conditions

$$p(\mathbf{x}_i, T_0) = p_0(\mathbf{x}_i), \quad (11)$$

$$[\nabla p(\mathbf{x}_i, t)] \cdot \hat{\mathbf{n}} = 0, \quad t \in (T_0, T_f], \quad (12)$$

where $\hat{\mathbf{n}}$ is a unit vector normal to the state-space boundary $\partial\mathcal{X}$. The zero-flux condition (12) prevents the agents from entering or leaving the state-space \mathcal{X} , such that the continuity equation assumptions are satisfied. Additionally, p must obey the normalization condition (9) and the state constraint

$$p(\mathbf{x}_i, t) = 0, \quad \mathbf{x}_i \notin \mathcal{X} \text{ and } t \in (T_0, T_f]. \quad (13)$$

Thus, the DOC problem consists of finding the optimal agent distribution p^* and microscopic controls \mathbf{u}_i^* that minimize the macroscopic cost J over the time interval $(T_0, T_f]$, subject to the parabolic PDE (10), the normalization condition (9), the initial and boundary conditions (11)–(12), and the state constraint (13).

DISTRIBUTED OPTIMAL CONTROL SOLUTION

The solution to the DOC problem can be obtained either via direct methods that discretize the DOC equations (9)–(13) with respect to space and time or via indirect methods that solve the optimality conditions numerically for the optimal agent PDF. Similarly to traditional feedback control, the

optimality conditions are derived by assuming that the agent microscopic control law is a function of the state, such as $\mathbf{u}_i = \mathbf{c}[p(t), t]$. Assume for simplicity that the random inputs \mathbf{w}_i and the diffusion term in (10) are both zero. Then, the advection-diffusion equation (10) reduces to the advection equation, and from the distributive property of the dot product and by a change of sign, it can be rewritten as the time-varying equality constraint

$$\frac{\partial p}{\partial t} + (\nabla p) \cdot \mathbf{f} + p(\nabla \cdot \mathbf{f}) = 0, \quad (14)$$

where function arguments are omitted for brevity.

Because (14) is a dynamic constraint that must be satisfied at all times, a time-varying Lagrange multiplier $\lambda = \lambda(\mathbf{x}_i, t)$ is used to adjoin the equality constraint (14) to the integral cost (7). Then, introducing the Hamiltonian

$$\mathcal{H} \equiv \mathcal{L} + \lambda[(\nabla p) \cdot \mathbf{f} + p(\nabla \cdot \mathbf{f})] = \mathcal{H}[p, \mathbf{u}_i, \lambda, t], \quad (15)$$

the DOC necessary conditions for optimality can be derived from the fundamental theorem of calculus of variations [33] such that the adjoint equation,

$$\dot{\lambda} = \frac{\partial \mathcal{H}}{\partial p} = \frac{\partial \mathcal{L}}{\partial p} + \lambda(\nabla \cdot \mathbf{f}), \quad (16)$$

and the optimality condition,

$$0 = \frac{\partial \mathcal{H}}{\partial \mathbf{u}_i} = \frac{\partial \mathcal{L}}{\partial \mathbf{u}_i} + \lambda \left[(\nabla p) \frac{\partial \mathbf{f}}{\partial \mathbf{u}_i} + p \frac{\partial}{\partial \mathbf{u}_i} (\nabla \cdot \mathbf{f}) \right], \quad (17)$$

are to be satisfied for $T_0 \leq t \leq T_f$, subject to the terminal condition

$$\int_{\mathcal{X}} \lambda(\mathbf{x}_i, T_f) d\mathbf{x}_i = - \left. \frac{\partial \phi}{\partial p} \right|_{t=T_f}. \quad (18)$$

A detailed proof of the above optimality conditions can be found in [19]. The optimality conditions for the DOC problem with nonzero random inputs can be found in [20].

The optimal agent distribution p^* is one that satisfies (16)–(18) along with the normalization condition (9), the initial and boundary conditions (11)–(12), and the state constraint (13). When all of these conditions are satisfied, the extremals can be tested using higher-order variations to verify that they lead to a minimum of the cost function. In particular, sufficient conditions for optimality can be obtained from the second-order derivatives of the Hamiltonian (15) with respect to \mathbf{u}_i , or a Hessian matrix that is positive definite for a convex Hamiltonian.

Numerical Solution

The DOC optimality conditions (16)–(18) consist of a set of parabolic PDEs for which analytical solutions are presently unknown. Therefore, indirect numerical methods of solution are needed to solve the optimality conditions numerically for the optimal agent PDF p^* , which can then be used to design the agent feedback control law $\mathbf{u}_i = \mathbf{c}[p^*(t), t]$, as shown in the next section. One indirect DOC method was

recently developed in [20] based on the generalized reduced gradient (GRG) method. The GRG method belongs to the class of nested analysis and design numerical methods for solving PDE-constrained optimization problems [34].

The GRG DOC method presented in [20] relies on iteratively solving the forward (14), adjoint (16), and optimality criterion (17) as a PDE-constrained minimization problem. Inspired by indirect methods for classical optimal control [23], in the GRG DOC method the microscopic control is parameterized as the sum of linearly independent Fourier basis functions and then used to obtain approximations of the costate and macroscopic state from the solution of the parabolic optimality conditions. Subsequently, holding the costate and macroscopic-state approximations fixed, the microscopic control input approximation is updated by a gradient-based algorithm that minimizes the augmented Lagrangian function and, ultimately, satisfies the third and final optimality condition (17).

The adjoint and forward equations are two coupled parabolic PDEs that can be solved efficiently using a modified Galerkin method, referred to as CINT [35], which is chosen for its nondissipative property [36], [37]. In CINT, the PDE solution is approximated by a linear combination of polynomial basis functions used to satisfy the PDE operator and Gaussian or radial basis functions used to enforce the boundary conditions at each grid point [35]. Given a control-law parameterization, the forward equation becomes a parabolic PDE with Neumann boundary conditions that can be solved numerically to obtain a macroscopic state approximation. Once an approximate solution is obtained from the forward equation, the adjoint equation becomes a parabolic PDE in λ with Dirichlet boundary conditions, and a numerical solution for λ can be obtained by a superposition of polynomial basis and radial basis functions [20]. The new and improved control input approximation is then held fixed and used to obtain new CINT approximations for p and λ . This iterative process is repeated until all three optimality conditions are satisfied within a user-specified tolerance, at which point the algorithm has converged to the optimal agent distribution p^* .

As in classical optimal control, direct DOC methods are typically easier to implement than indirect methods. A direct NLP method of solution is reviewed here and analyzed in detail in [18]. The agent PDF p is a conserved quantity that satisfies the Hamilton equations [18]. As a result, the DOC problem can be discretized using a conservative finite-volume (FV) numerical scheme for representing and evaluating PDEs in algebraic form. This method does not incur dissipative errors associated with coarse-grained state discretization, a method that may require reduction in computation [38]. Using the FV approximation, the continuous DOC problem is discretized about a finite set of collocation points and then transcribes into a finite-dimensional NLP that can be solved using an efficient mathematical programming algorithm, such as SQP.

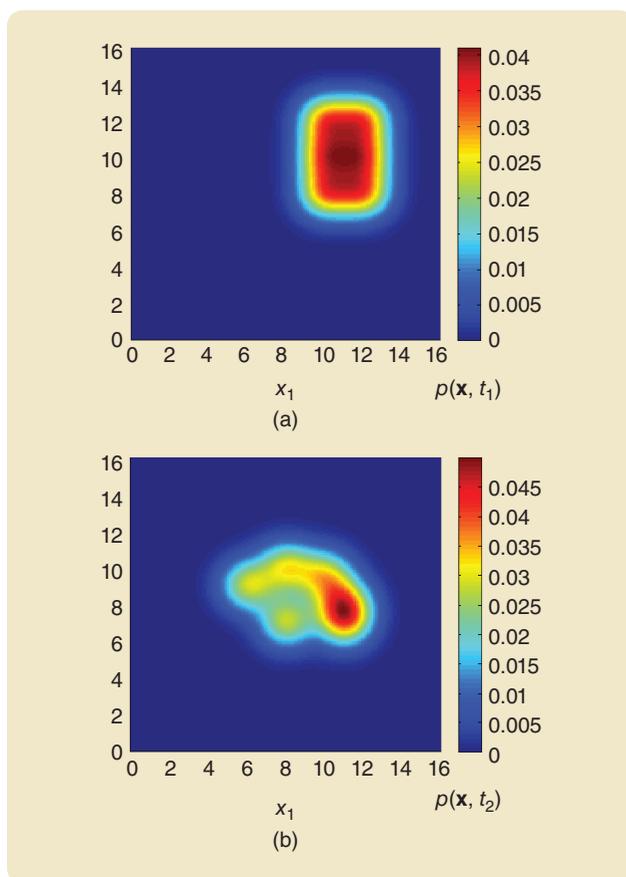


FIGURE 1 An example of a probability density function (PDF) modeled by a six-component Gaussian mixture. (a) An example of a PDF plotted at time t_1 and (b) an example of PDF plotted at time t_2 , for a state vector $\mathbf{x} = [x_1 \ x_2]^T$.

To discretize the DOC problem, the agent PDF p must be parameterized over the solution domain \mathcal{X} . As a simple example, a finite Gaussian-mixture model is chosen here to provide this parametric approximation as a superposition of z components with Gaussian PDFs, denoted by f_1, \dots, f_z , and corresponding mixing proportions or weights, denoted by w_1, \dots, w_z . The n -dimensional multivariate Gaussian PDF

$$f_j(\mathbf{x}_i, t) = \frac{1}{(2\pi)^{n/2} |\Sigma_j|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_j)} \quad (19)$$

is referred to as the *component density* of the mixture and is characterized by a time-varying mean vector $\boldsymbol{\mu}_j \in \mathbb{R}^n$ and a time-varying covariance matrix $\Sigma_j \in \mathbb{R}^{n \times n}$ with $j = 1, \dots, z$ [39]. Assume that at any $t \in (T_0, T_f]$ the optimal agent distribution can be represented as

$$p(\mathbf{x}_i, t) = \sum_{j=1}^z w_j(t) f_j(\mathbf{x}_i, t), \quad (20)$$

where $0 \leq w_j \leq 1$ for all j , $\sum_{j=1}^z w_j = 1$, and z is fixed and chosen by the user. Then, an optimal (time-varying) agent distribution p^* can be obtained by determining the optimal trajectories of the mixture model parameters from the DOC problem. An example of time-varying distribution

described by a three-component Gaussian mixture with a two-dimensional state vector $\mathbf{x} = [x_1 \ x_2]^T$ is shown in Figure 1 at two moments in time t_1 and t_2 , where, in this case, $z = 6$, and the mixture centers $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_6$ and weights w_1, \dots, w_6 all change over time.

The mixture model parameters to be optimized over time are the weights w_j , the elements of $\boldsymbol{\mu}_j$, and the variances and covariances in $\boldsymbol{\Sigma}_j$, with $j = 1, \dots, z$. In addition to satisfying the DOC constraints and optimality conditions, the mixture model parameters must be determined such that the component densities f_1, \dots, f_z are nonnegative and obey the normalization condition for all $t \in (T_0, T_f]$. This is accomplished by discretizing the continuous DOC problem in state space and time, about a finite set of collocation points in $\mathcal{X} \times (T_0, T_f]$. Let Δt denote a constant discretization time interval and k denote the discrete time index, such that $\Delta t = (T_f - T_0)/K$, and thus $t_k = k\Delta t$, for $k = 0, \dots, K$. It is assumed that the microscopic control inputs, \mathbf{u}_i , with $i = 1, \dots, N$, are piecewise constant during every time interval, and that the agent distribution at time t_k can be represented by

$$p_k = p(\mathbf{x}_i, t_k) = \sum_{j=1}^z w_j(t_k) f_j(\mathbf{x}, t_k) \\ \equiv \sum_{j=1}^z w_{jk} \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}_{jk}|^{1/2}} e^{[-(1/2)(\mathbf{x}_i - \boldsymbol{\mu}_{jk})^T \boldsymbol{\Sigma}_{jk}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_{jk})]}. \quad (21)$$

The set of weights $\{w_{jk}\}$ and the elements of $\boldsymbol{\mu}_{jk}$ and $\boldsymbol{\Sigma}_{jk}$, for all j, k , are grouped into a vector $\boldsymbol{\chi}$ that represents the trajectories of the mixture model parameters in discrete time.

The FV approach partitions the state-space \mathcal{X} into FVs defined by a constant discretization interval $\Delta \mathbf{x} \in \mathbb{R}^n$ that are each centered about a collocation point $\mathbf{x}_l \in \mathcal{X} \subset \mathbb{R}^n$, with $l = 1, \dots, L$ [38]. Let $p_{l,k}$ and $\mathbf{u}_{l,k}$ denote the finite-difference approximations of $p(\mathbf{x}_l, t_k)$ and $\mathbf{c}[p(\mathbf{x}_l, t_k)]$, respectively. Then, the finite-difference approximation of the advection equation is obtained by applying the divergence theorem to (14) for every FV such that $p_{k+1} = p_k + \Delta t \rho_k$, where

$$\rho_k \triangleq - \int_S [p_k \mathbf{f}(p_{l,k}, \mathbf{u}_{l,k}, t_k)] \cdot \hat{\mathbf{n}} \, dS, \quad (22)$$

and S and $\hat{\mathbf{n}}$ denote the FV boundary and unit normal, respectively.

Now, letting $\Delta \mathbf{x}_{(j)}$ denote the j th element of $\Delta \mathbf{x}$, the discretized DOC problem can be written as the finite-dimensional NLP

$$\min J_D = \sum_{j=1}^n \Delta \mathbf{x}_{(j)} \sum_{l=1}^L [\phi_{l,K} + \Delta t \sum_{k=1}^K \mathcal{L}(p_{l,k}, \mathbf{u}_{l,k}, t_k)], \\ \text{subject to } p_{k+1} - p_k - \Delta t \rho_k = 0, \quad k = 1, \dots, K, \\ \sum_{j=1}^n \Delta \mathbf{x}_{(j)} \sum_{l=1}^L p_{l,k} - 1 = 0, \quad k = 1, \dots, K, \\ p_{l,0} = p_0(\mathbf{x}_l), \quad \mathbf{x}_l \in \mathcal{X}, \\ p_{l,k} = 0, \quad \mathbf{x}_l \in \partial \mathcal{X}, \quad k = 1, \dots, K, \quad (23)$$

where $\phi_{l,k} = \phi(p_{l,k})$ is the terminal constraint. The solution $\boldsymbol{\chi}^*$ of the above NLP can be obtained using an SQP algo-

rithm that solves the KKT optimality conditions by representing (23) as a sequence of unconstrained quadratic programming subproblems [40]. The reader is referred to [18] for a detailed description and analysis of the above direct DOC method, including the computational complexity analysis and comparison with classical direct optimal control methods.

Distributed Optimal Control Feedback Control Law

Although the optimal agent control laws are obtained together with the optimal agent PDF, it is often more useful in practice to separate the planning and control designs such that state measurements can be fed back and accounted for by the agent while attempting to realize or “track” the optimal time-varying PDF p^* . By this approach, the actual agent PDF can be estimated and considered by the feedback control law, and the feedback control can be designed to achieve additional local constraints or account for communication constraints. Various techniques, including Voronoi diagrams, Delaunay triangulations, and potential field methods, have been proposed in the literature to design agent control laws that achieve a desired PDF evolution over time [2], [5], [14]. Because in DOC the optimal PDF is computed subject to the agent dynamics (6), the agents are guaranteed to be able to follow p^* over the time interval $(T_0, T_f]$. The agent control law can also be designed to provide additional guarantees, such as closed-loop stability and mutual collision avoidance, accounting for communication protocols for exchanging state information.

In this tutorial, the feedback control law design is illustrated by means of a potential field method combined with a decentralized kernel density estimation algorithm that requires the network of N agents to be connected at any time $t \in (T_0, T_f]$. Although p^* can be determined offline by solving the optimality conditions in the section “Numerical Solution” via centralized or decentralized optimization algorithms, once the agents are deployed in an uncertain environment, their individual microscopic state (or output) measurements must be considered to ensure that they follow p^* over time. Thus, to track p^* , the microscopic feedback control law $\mathbf{c}[\cdot]$ must be a function of the deviations between p^* and the actual agent PDF obtained from the microscopic state \mathbf{x}_i , with $i = 1, \dots, N$. Assume for simplicity that the microscopic state \mathbf{x}_i is fully observable for all i , and let $\hat{\mathbf{x}}_i(t)$ denote its estimated value at time t . Then, estimating the actual agent PDF at time t requires N observations $\hat{\mathbf{x}}_1(t), \dots, \hat{\mathbf{x}}_N(t)$.

In large systems of agents, it is unrealistic to assume that each agent can communicate and acquire the state estimate of all other $(N - 1)$ agents at all times. To implement the DOC approach in a decentralized network, the actual agent density must be approximated locally by the i th agent, without requiring direct communication with all other agents in the network. This can be achieved through a decentralized adaptation of the nonparametric technique

known as kernel density estimation (KDE). In KDE, each node of a decentralized network repeatedly exchanges data with its neighbors through information spreading and then performs a local KDE calculation. Through this process, each local estimate separately converges asymptotically to the distribution that can be obtained using the centralized KDE method. Other decentralized techniques have been presented for estimating a distribution from a data set using distributed expectation maximization (EM) algorithms. While other EM techniques may perform poorly with small data sets, remain trapped in local maxima, or display sensitivity to initial parameter choices, the distributed KDE approach described in this section was found to not suffer from any of these limitations.

KDE is a well-known nonparametric approach for estimating a PDF from a set of independent and identically distributed (iid) data samples. Although in DOC the agents are not independent, their state values can be assumed to be iid for the purpose of estimating their instantaneous PDF without significant loss of accuracy. Given the data set $\{\hat{\mathbf{x}}_i(t): \hat{\mathbf{x}}_i \in \mathbb{R}^n, t \in (T_0, T_f), i = 1, \dots, N\}$, assumed drawn from an unknown PDF, $p(\mathbf{x}_i, t)$, the kernel density estimation takes the form

$$\hat{p}(\mathbf{x}_i, t) = \sum_{i=1}^N \gamma_i \mathbf{K}_{\mathbf{H}_i}[\mathbf{x} - \hat{\mathbf{x}}_i(t)], \quad \mathbf{x} \in \mathbb{R}^n, \quad (24)$$

where $\gamma_1, \dots, \gamma_N$ are the weighting coefficients that satisfy the condition $\sum_{i=1}^N \gamma_i = 1$ [41]. The i th kernel centered at $\hat{\mathbf{x}}_i(t)$ is defined as

$$\mathbf{K}_{\mathbf{H}_i} = |\mathbf{H}_i|^{-1/2} \mathbf{K}\{\mathbf{H}_i^{-1/2}[\mathbf{x} - \hat{\mathbf{x}}_i(t)]\}, \quad (25)$$

where the kernel function \mathbf{K} is a user-defined n -variate nonnegative symmetric real-valued function [42]. The bandwidth matrix \mathbf{H}_i is a parameter that controls the smoothing of the KDE algorithm, and it must be positive definite and symmetric. With appropriate parameter choices, KDE has been shown to be an effective method for estimating the underlying PDF, often requiring a few samples to produce adequate results [42].

A decentralized KDE (DKDE) algorithm that does not require centralized processing and is asymptotically consistent with the centralized version for fully connected networks was recently proposed in [43]. This DKDE algorithm uses an information-sharing protocol to incrementally exchange kernel information between agents with a bounded communication radius r , until a complete and accurate approximation of the global KDE is achieved by each agent in the network. It has been shown in [44] that, for a fully connected network, the connectivity structure will only affect the convergence speed and will not worsen the estimation accuracy. In DKDE, each agent maintains a local estimate of the agent PDF, governed by a stored kernel set, $\mathbf{S}_i = \{\langle \gamma_{i,k}, \hat{\mathbf{x}}_{i,k}, \mathbf{H}_{i,k} \rangle, k = 1, \dots, N_i\}$, where $\hat{\mathbf{x}}_{i,k}$ denotes the position of agent k perceived by agent i , N_i is the number

of kernels stored by agent i , and $\mathbf{H}_{i,k}$ and $\gamma_{i,k}$ are the bandwidth matrix and weighting coefficient of the k th kernel stored by agent i . Initially, the kernel set of each agent only contains the kernel generated using its own position. Each agent i also maintains a neighbor set, defined as agents located within a distance r of \mathbf{x}_i . Then, using an information spreading protocol, the agent can choose to communicate with a neighbor in the set at random and mutually compare kernel sets. When an agent sees a newer or previously unknown kernel, it saves the corresponding information into its own stored kernel set. Then, a new neighbor is chosen at random, and the process is repeated over time.

The information exchanged by the agents should include their state estimates, kernel parameters, indices, and time stamps to enable overwriting of old data. For homogeneous networks, such as those considered in this tutorial, the bandwidth matrices $\mathbf{H}_{i,k}$ and weighting parameters $\gamma_{i,k}$ may be defined to be consistent across the network, making their communication unnecessary and reducing communication requirements. For example, in the simplest case, the bandwidth matrix can be chosen as $\mathbf{H}_{i,k} = c\mathbf{I}_2$, where c is a constant and \mathbf{I}_2 is the two-dimensional identity matrix, and $\gamma_{i,k} = 1/N_i$ for all i and k . Then, from a stored set of agent state estimates, each agent can generate the corresponding kernels and combine them to obtain a local estimate of the actual agent PDF. Consider the standard two-dimensional Gaussian kernel function

$$\mathbf{K}(\mathbf{x}) = (2\pi)^{-1} e^{-1/2 \mathbf{x}^T \mathbf{x}}, \quad \mathbf{x} \in \mathbb{R}^n, \quad (26)$$

which is consistent with the chosen parameterization of the optimal agent PDF. The kernels of agent i are

$$\mathbf{K}_{\mathbf{H}_{i,k}} = |\mathbf{H}_{i,k}|^{-1/2} \mathbf{K}\{\mathbf{H}_{i,k}^{-1/2}[\mathbf{x} - \hat{\mathbf{x}}_i(t)]\}, \quad (27)$$

and its estimate of the agent PDF is

$$\hat{p}_i(\mathbf{x}_i, t) = \sum_{k=1}^{N_i} \gamma_{i,k} \mathbf{K}_{\mathbf{H}_{i,k}}[\mathbf{x} - \hat{\mathbf{x}}_i(t)]. \quad (28)$$

Once agent i has performed its local density estimation, the feedback control law is computed from the optimal agent PDF p^* using a potential field approach.

Potential field methods are commonly used for obstacle avoidance by generating a virtual navigation function that “repels” the robot or vehicle away from obstacles and “pulls” it toward a desired state or configuration [45]. Stabilization and potential navigation methods for nonholonomic systems can also be obtained with some additional precautions, such as the use of time-varying smooth control laws [46]–[48], discontinuous feedback control [49], and switched control systems [50], [51]. A potential field method for designing DOC feedback control laws that attract or pull the agents toward the optimal PDF p^* was recently presented in [19]. By this approach, since p^* represents the goal PDF of all N agents, the event of an agent assuming a state value $\hat{\mathbf{x}}_i$ downgrades the probability mass such that the

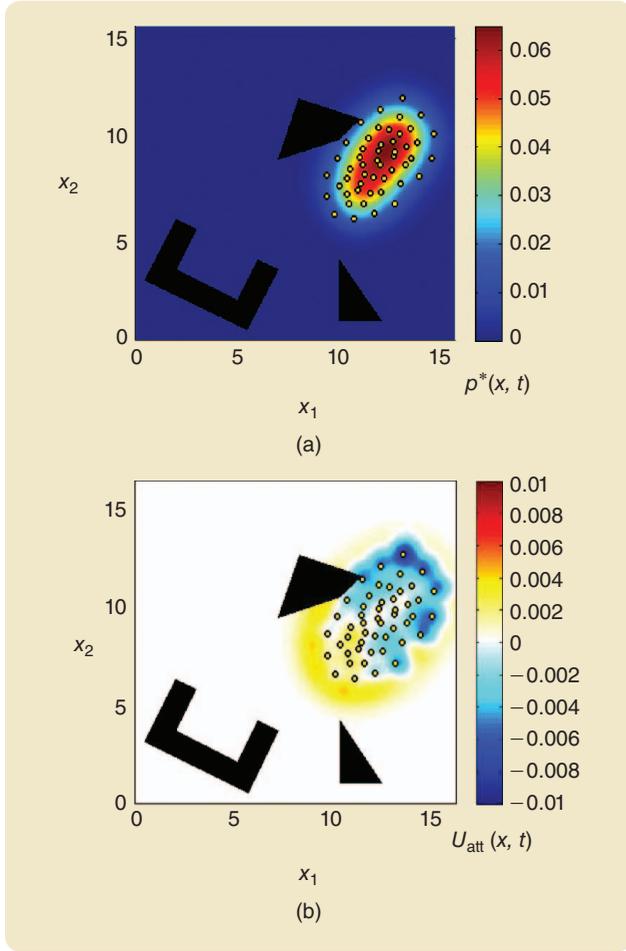


FIGURE 2 A network of agents navigating an obstacle-populated environment (the obstacles are shown in black). (a) An optimal probability density function and (b) attractive potential for the network of agents.

probability of another agent in the network assuming the same state value is decreased.

Thus, agent i must construct a feedback control law using its knowledge of p^* and the estimated agent distribution \hat{p}_i obtained via DKDE as follows. Let the attractive potential of agent i be defined as

$$U_{\text{att}}^i(\mathbf{x}_i, t) \triangleq \frac{1}{2} [\hat{p}_i(\mathbf{x}_i, t + \delta t) - p^*(\mathbf{x}_i, t + \delta t)]^2, \quad (29)$$

where δt is a time-shift parameter that allows the control law to look ahead in time to prevent agents from lagging behind. The estimate $\hat{p}_i(\mathbf{x}_i, t + \delta t)$ is computed by stepping the advection equation (14) forward in time by δt from the DKDE estimate $\hat{p}_i(\mathbf{x}_i, t)$ in (28). An example of attractive potential and the corresponding optimal agent distribution is shown in Figure 2, where the agents shown by yellow circles and are navigating an environment populated with obstacles. Then, the potential navigation function for agent i can be generated as the sum of the attractive potential in (29) and a repulsive potential U_{rep}^i constructed to avoid mutual collisions or obstacles

$$U_i(\mathbf{x}_i, t) = w_a U_{\text{att}}^i(\mathbf{x}_i, t) + w_r U_{\text{rep}}^i(\mathbf{x}_i, t), \quad (30)$$

where w_a and w_r are user-defined weighting coefficients.

A feedback control law that follows the navigation function (30) can be obtained from its gradient

$$\mathbf{u}_i = [v_c \ Q(\hat{\theta}_i, -\nabla U_i)]^T, \quad (31)$$

where the minimum difference between the desired heading angle $\theta(-\nabla U_i)$ and the agents' actual heading angle $\hat{\theta}_i$ is

$$Q(\cdot) = \{a(\hat{\theta}_i) - a[\theta(-\nabla U_i)]\} \text{sgn}\{a[\theta(-\nabla U_i)] - a(\hat{\theta}_i)\}, \quad (32)$$

where $\text{sgn}(\cdot)$ is the sign function, $a(\cdot)$ is an angle wrapping function, and v_c is the agent speed.

The DOC concept is illustrated in Figure 3. As a first step, the microscopic agent states are mapped to a macroscopic description via the restriction operator p , defined based on the desired system performance or cost function. Subsequently, the optimal evolution of restriction operator p^* is determined with significant computational savings using a direct or indirect DOC numerical method. Subsequently, the optimal evolution of restriction operator p^* is determined with significant computational savings using a direct or indirect DOC numerical method. Subsequently, microscopic agent control laws that meet the desired optimal performance are obtained, for example using a potential field method.

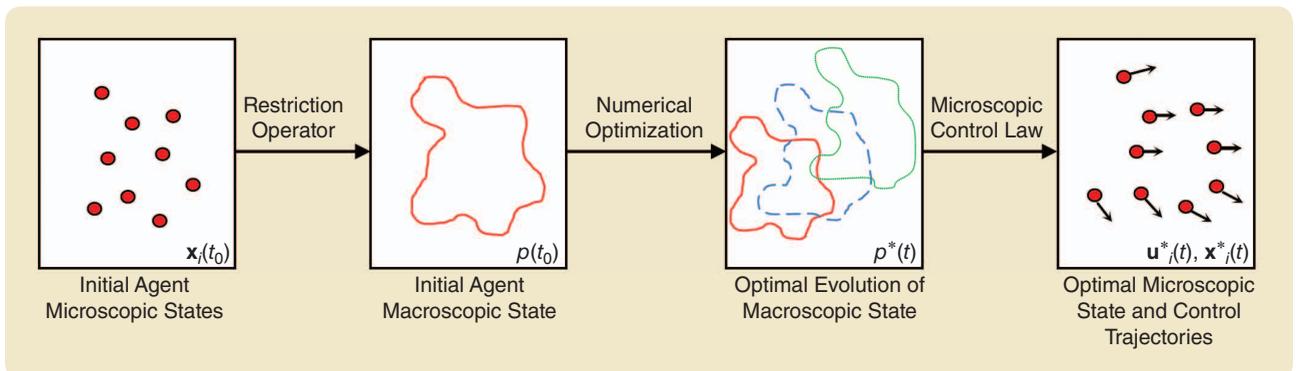


FIGURE 3 The conceptual steps of the distributed optimal control approach.

APPLICATION EXAMPLES

Although in principle, it is applicable to other restriction operators, the DOC approach is described in this tutorial using PDFs because they lend themselves to a broad range of applications involving networks or teams of collaborative agents, such as sensors, robots, or autonomous vehicles. Also, because DOC optimizes a general functional of the macroscopic state and not the expectation like many distributed control methods [3], [11], [16], it is applicable to a broad range of objectives that may include, for example, the use of information-theoretic functions, as shown in the next section.

Very-Large-Scale Robotic Systems Path Planning

A common application of VLRS is to safely and efficiently navigate a complex environment while performing multiple tasks, such as searching for a target or communicating with a central station. These environments may contain hazards to be avoided such as obstacles or water bodies, as well as a destination that the agents must reach with minimum energy consumption. Other tasks may require formation maintenance during navigation, for example to provide coverage or maintain connectivity in the network of agents. This class of problems, commonly known as multiagent path planning, is known to be computationally very burdensome for large numbers of agents. In particular, it was recently shown that optimizing the trajectories of N agents in an obstacle-populated environment is PSPACE-hard and, thus, is generally considered a computationally intractable problem for large N because it requires exponential deterministic time in the worst case [52].

This example considers a system of N collaborative unicycle robots traveling through an obstacle-populated compact space $\mathcal{W} \subset \mathbb{R}^2$, referred to as the *workspace*, and occupied by M obstacles $\mathcal{B}_1, \dots, \mathcal{B}_M$, where $\mathcal{B}_j \subset \mathcal{W}$. The dynamics of each agent are described by the nonlinear unicycle model

$$\dot{x}_i = v_i \cos \theta_i, \quad \dot{y}_i = v_i \sin \theta_i, \quad \dot{\theta}_i = \omega_i, \quad (33)$$

where $\mathbf{x}_i = [x_i \ y_i \ \theta_i]^T$ is the microscopic state of agent i , x_i and y_i are the agent's xy -coordinates, θ_i is the heading angle, and $i = 1, \dots, N$. The microscopic control vector is $\mathbf{u}_i = [v_i \ \omega_i]^T$, where v_i and ω_i are the linear and angular velocities, respectively. Given an initial distribution $p_0(\mathbf{x})$, the agents must travel in \mathcal{W} to meet a goal distribution $g(\mathbf{x})$, while avoiding obstacles, maintaining a triangular

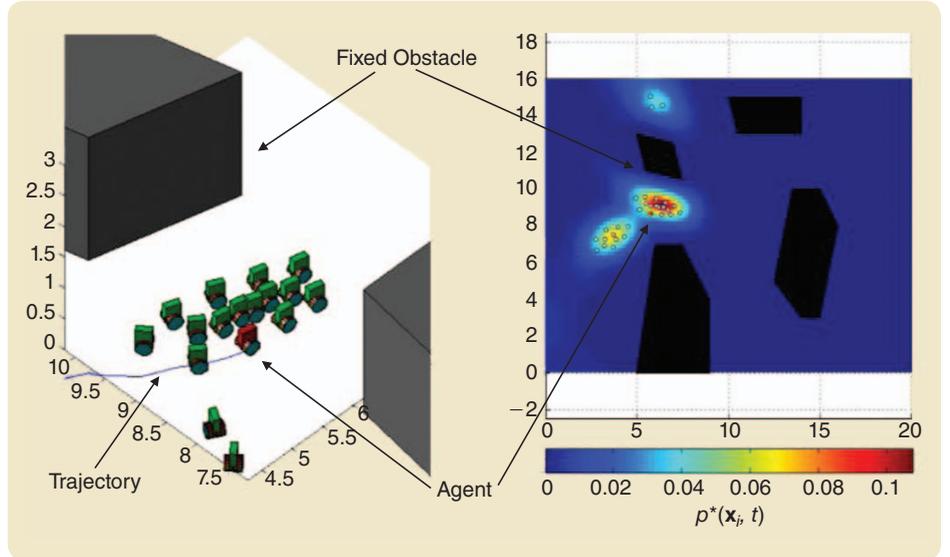


FIGURE 4 An example representation of robots as microscopic agents following an optimal macroscopic description.

formation, and minimizing energy consumption. The goal distribution p_0 and all M obstacles are assumed fixed and known a priori.

Figure 4 shows both the microscopic and macroscopic view of the same network of unicycle robots controlled using an optimal PDF obtained via DOC, as they enter a narrow passage. A red agent is identified in both views, and its trajectory is plotted as a blue line for illustration. It can be seen that, at this stage in the simulation, two robots from the lower group of agents are about to join the denser group above, in this case to be able to navigate around the bottom obstacle.

The agent objective of reaching the goal distribution g can be formulated using the instantaneous Kullback-Leibler (KL) divergence at time t

$$D(p \parallel g) = \int_{\mathcal{X}} p(\mathbf{x}_i, t) \log_2 \frac{p(\mathbf{x}_i, t)}{g(\mathbf{x}_i)} d\mathbf{x}_i, \quad (34)$$

where, by definition, the support set of p is contained by the support set of g , and the value $0 \log_2(0/0)$ is replaced with zero for continuity [53]. Although the KL divergence is not a true distance function because it is not symmetric and does not obey the triangle inequality, it is a suitable objective function because its value increases when the difference between p and g increases, and vice versa. Also, the KL divergence of p and g is zero when the two distributions are equal.

The agent objective of avoiding collisions with known obstacles can be formulated by first generating a repulsive potential U_{rep} based on the obstacle geometries and locations in \mathcal{W} , as shown in [45], and then minimizing the product pU_{rep} . Additional objectives, such as holding an equilateral triangular distribution pattern, can also be

introduced by specifying higher-order moments of the distribution. For example, a triangular formation can be specified by imposing a constant distance a , specified by the user, between the centers of a Gaussian mixture with ($z = 3$) components. The center of the j th component is the first moment or mean vector

$$\boldsymbol{\mu}_j(t) = \int_{\mathcal{W}} \mathbf{x}_i f_j(\mathbf{x}_i, t) d\mathbf{x}_i, \quad (35)$$

thus, the desired formation can be maintained by minimizing the objective function

$$S(p) = \sum_{j \neq l} \|\boldsymbol{\mu}_j(t) - \boldsymbol{\mu}_l(t)\| - a, \quad j, l = 1, \dots, z. \quad (36)$$

Modeling the agent energy consumption as a quadratic function of the microscopic control, the DOC cost function to be minimized is

$$J = \int_{T_0}^{T_f} [w_d D(p \| g) + w_s S(p) + \int_{\mathcal{W}} (w_r p U_{\text{rep}} + w_e \mathbf{u}_i^T \mathbf{R} \mathbf{u}_i) d\mathbf{x}_i] dt, \quad (37)$$

where \mathbf{R} is a diagonal positive-definite matrix, and the scalar weights w_d , w_r , w_s , and w_e are chosen by the user based on the desired tradeoff between the four competing objectives.

As described in the section ‘‘Distributed Optimal Control Solution,’’ once an optimal agent distribution p^* is obtained from the DOC problem (33)–(39), the microscopic control laws are obtained using the gradient of the potential function, as shown in (31). In this example, the trajectories of $N = 300$ unicycle agents are simulated via DOC in a workspace with one obstacle (Figure 5), with initial and goal distributions, p_0 and g , plotted in Figure 5, over the time interval (0; 22] h. The initial microscopic states \mathbf{x}_{i_0} are obtained by sampling p_0 . The cost function weights are $w_d = 15$, $w_q = 100$, $w_r = 0.15$, and $w_e = 1.5$, and $z = 3$. Time is discretized in intervals of $\Delta t = 1$ h, such that $K = 22$, and the state space is discretized using $X = 900$ FVs. The optimal agent distribution and the microscopic state values obtained by implementing the microscopic control law in (31) are plotted in Figure 6 at four sample moments in time: (a) $t = 0$ h, (b) $t = 8$ h, (c) $t = 15$ h, and (d) $t = 21$ h. It can be seen that the agents move from an initial distribution that does not obey the desired formation [Figure 6(a)] to an equilateral formation that also is able to avoid the obstacle and reach the goal distribution at the final time.

Control of Multiple Agent Distributions

Many distributed systems have different classes of agents that must interact to achieve collaborative objectives. The DOC methodology can be applied to such systems by representing the distribution of each class of agents as a distinct PDF. This concept is illustrated here by considering a toy problem in which three classes of agents are to be controlled subject to the same microscopic dynamics to match three final distributions defined by the color of a target image. In this example, the agents can be thought of as masses of color pigments that coordinate to form the desired distribution of pixels in the given image. Although the approach can be easily extended to classes with different dynamic equations, for simplicity in this example all agents are modeled by single integrator dynamics with an additive Gaussian noise term

$$\dot{\mathbf{x}}_i(t) = \mathbf{u}_i(t) + \sigma \mathbf{I}_2 \mathbf{w}_i(t), \quad (38)$$

where the agent microscopic state is $\mathbf{x}_i = [x_i \ y_i]^T \in \mathcal{W}$, $\mathbf{u} = [u_x \ u_y]^T$ is the control, $\mathbf{w} \in \mathbb{R}^2$ is a vector of iid Gaussian random disturbances, and σ is a constant coefficient.

The performance of the system is represented in terms of three agent PDFs, $p_R(\mathbf{x}_i, t)$, $p_G(\mathbf{x}_i, t)$, and $p_B(\mathbf{x}_i, t)$, corresponding to three classes of agents denoted by red (R), green (G), and blue (B), respectively. Because the goal distributions, denoted by $Q_R(\mathbf{x}_i)$, $Q_G(\mathbf{x}_i)$, and $Q_B(\mathbf{x}_i)$, occupy the same workspace \mathcal{W} , the RGB color model is adopted to show the superposition of the three densities as one target

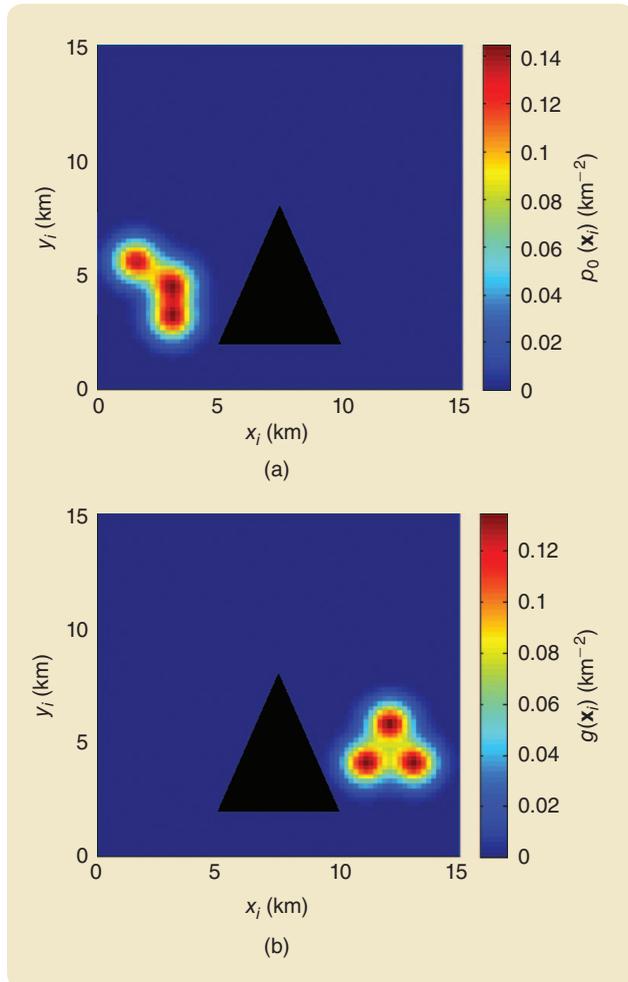


FIGURE 5 Agent distributions for a workspace with one obstacle (solid black). (a) Initial agent distribution and (b) goal agent distribution.

This article describes a coarse-grained optimal control approach for multiscale dynamical systems, referred to as distributed optimal control.

image in \mathcal{W} . The target image consists of $m \times n$ pixels and is decomposed into its RGB color intensities, such that the intensity of each color is defined uniquely in \mathcal{W} . Let the goal distribution $Q_I(\mathbf{x}_i)$ denote the intensity of color $I \in C$, where $C = \{R, G, B\}$ is the color index set. Because, in this case, the RGB color intensities will differ depending on the target image, the agent distributions will not technically be PDFs but rather conserved densities. To guarantee that the agent densities are conserved, the agent densities are defined such that $\int_{\mathcal{W}} p_i(\mathbf{x}_i, t) d\mathbf{x}_i = \int_{\mathcal{W}} Q_I(\mathbf{x}_i) d\mathbf{x}_i$ for all t, I , where typically $\int_{\mathcal{W}} Q_I(\mathbf{x}_i) d\mathbf{x}_i \neq 1$.

The cost function to be minimized subject to the agent dynamics (38) represents the sum of the energy expenditures and the errors between the final agent densities and the corresponding color intensity in the target image

$$J = \sum_{I \in C} \int_{\mathcal{W}} \left\{ [p_I(\mathbf{x}_i, T_f) - Q_I(\mathbf{x}_i)] + \int_{T_0}^{T_f} \mathbf{u}_I^T \mathbf{R}_I \mathbf{u}_I dt \right\} d\mathbf{x}_i, \quad (39)$$

where the weight matrices \mathbf{R}_I are chosen by the user. Then, given the initial RGB distribution in Figure 7(a), the agents must navigate to the target RGB image in Figure 7(b).

Because the agent dynamics in (38) are characterized by random disturbances, the DOC problem takes the form of (6)–(10) and, thus, it is solved using the GRG indirect DOC numerical method described in the section “Numerical Solution” and presented in [20]. By this approach, the three agent PDFs are optimized over time, and the final distribution obtained at time T_f is shown in Figure 7(c). As another example, given an initial uniform RGB distribution analogous to that in Figure 7(a) (but with different intensities)

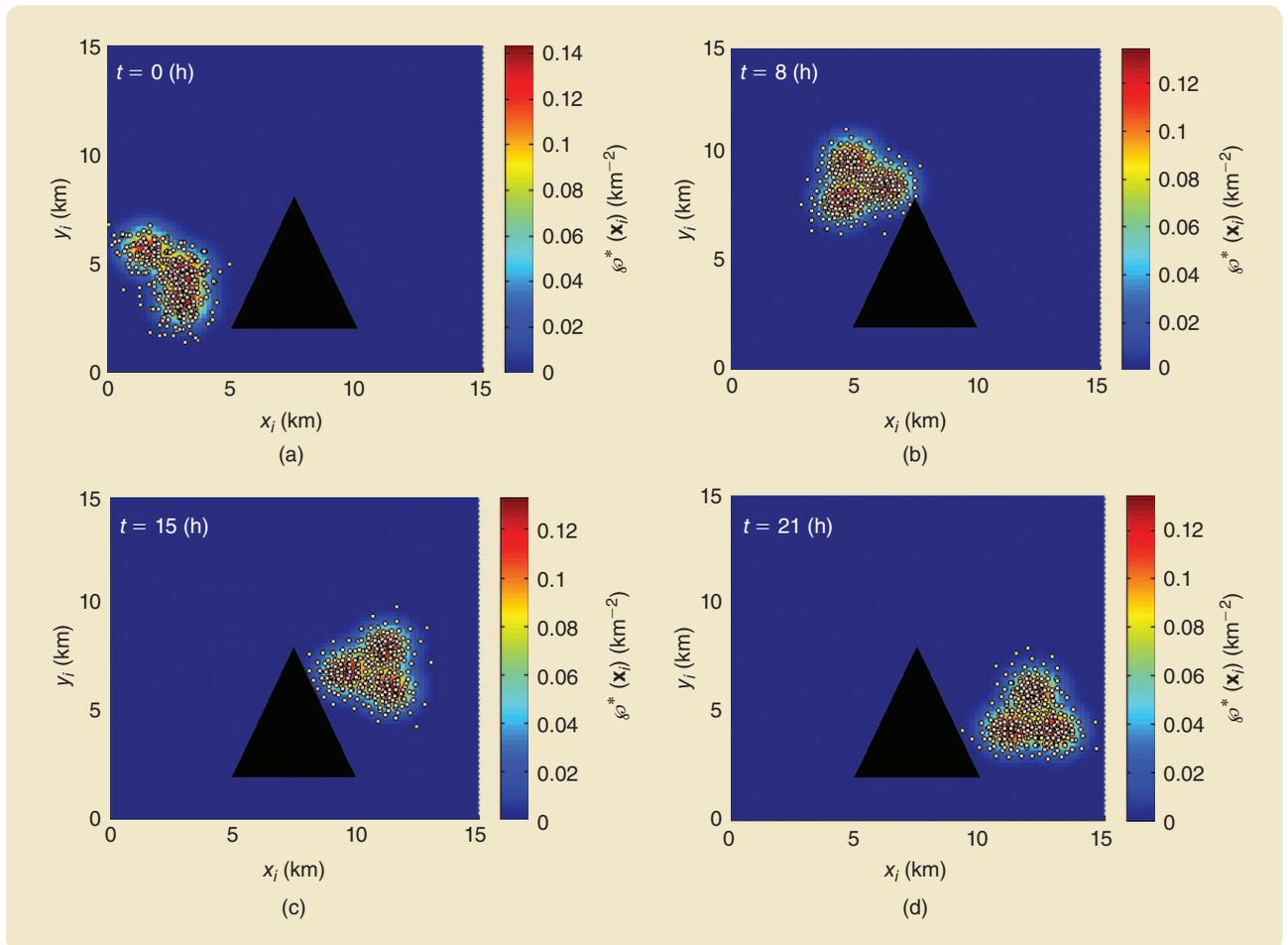


FIGURE 6 The optimal evolution of agent distribution and microscopic state (white circles) for $N = 300$ microscopic agents at four sample instants in time. (a) $t = 0$ (h), (b) $t = 8$ (h), (c) $t = 15$ (h), and (d) $t = 21$ (h).

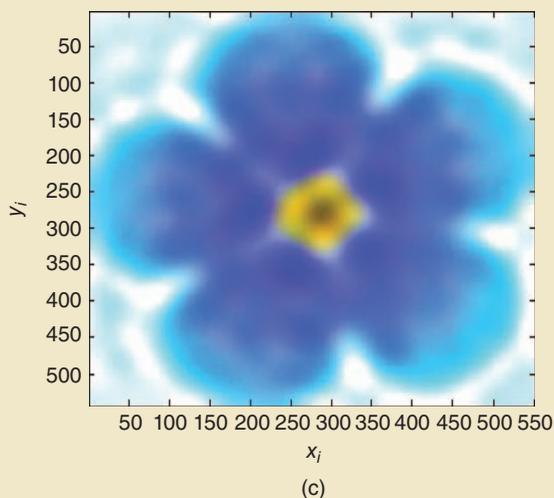
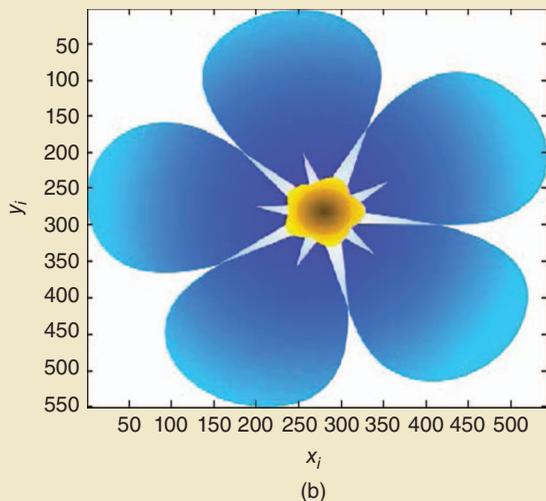
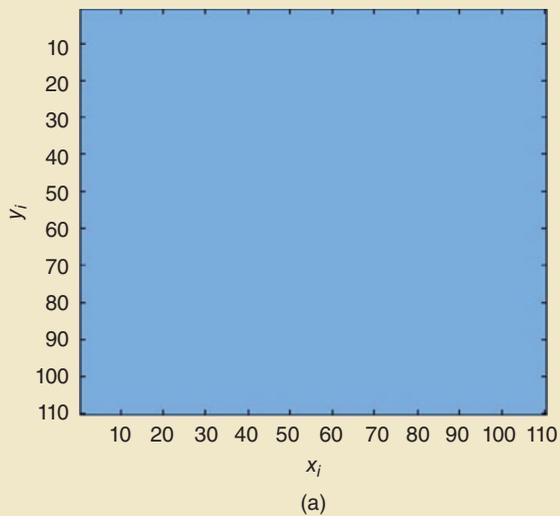


FIGURE 7 Agent distributions for a multiple-distribution distributed optimal control problem. (a) Initial agent distributions, (b) target red-green-blue image, and (c) final agent distributions.

and the target image shown in Figure 8(a), the GRG DOC solution achieves the final RGB distribution shown in Figure 8(b), subject to the agent dynamics in (38). The PDFs evolutions and individual agents are omitted here for simplicity, but it can be seen from Figures 7 and 8 that the control objective of matching the target images is well accomplished by the DOC method, while also minimizing energy. In these examples, the final agent distributions appear blurry because of the diffusion effect of the random disturbances w_i .

CONCLUSIONS AND RECOMMENDATIONS

This article provides an introduction to DOC methods for planning and control in multiscale dynamical systems. The DOC approach seeks to extend the capabilities of classical optimal control to multiscale dynamical systems of interacting agents in which optimal state and control laws must be determined for each (microscopic) agent but the cost function depends on the state of all agents over time. DOC methods rely on a restriction operator, such as a time-varying PDF or its moments, to reduce the computation required by optimizing system dynamics over large temporal and/or spatial scales. Because the resulting closed-loop

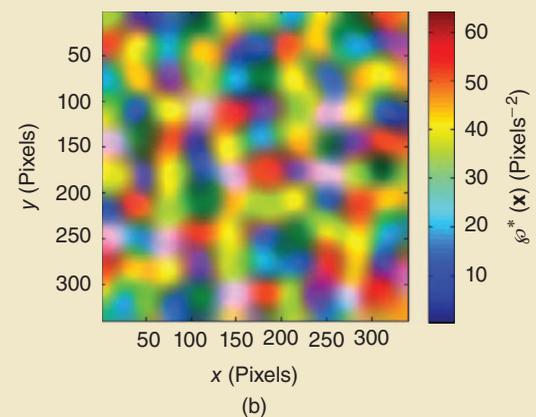
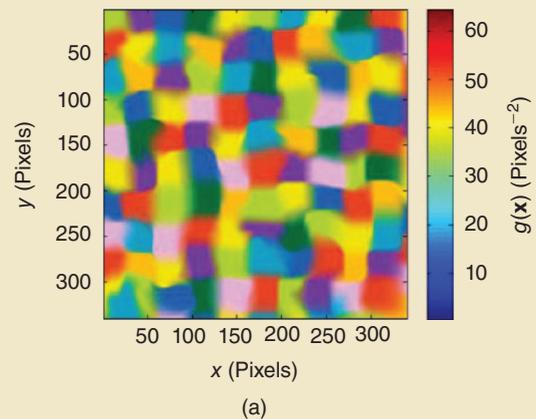


FIGURE 8 Images for a second multiple-distribution distributed optimal control problem. (a) Target red-green-blue image and (b) final agent distribution.

This tutorial provides an introductory overview of the DOC problem and compares the DOC to existing optimal control methods.

dynamics can be shown to be a Hamiltonian system, the agent performance is not compromised by the discretization carried out by direct DOC methods.

The agent performance can, however, be improved substantially by DOC methods that afford a general form of the PDF. By constraining the structure of the PDF, for example by parameterizing the PDF using a relatively small number of Gaussian mixture components, the computation required is reduced at the expense of optimality. Furthermore, a PDF operator may not be well suited to problems where it is important to impose highly accurate hard constraints on the PDF evolution, such as the boundaries in the petals and colors of Figure 7(b). Once the optimal PDF is determined, the agent behavior obtained by the proposed feedback control laws is found to be robust. Because the control law of each agent is based on the optimal PDF, it does not require simulating the advection equation and is computationally inexpensive. However, a control design different from the one reviewed in this article may be needed when agents must meet local performance objectives in addition to the system-level objectives expressed by the DOC cost function.

To date, DOC methods have been developed and demonstrated for trajectory optimization and control in MAS comprised of many dynamical systems with decoupled dynamics but mutual objectives. Future research will address multiscale systems with coupled agent dynamics as well as new restriction operators, such as maximum likelihood density estimators. To implement the DOC methods reviewed in this tutorial, the multiscale dynamical system should be composed of distributions of agents with the same dynamic characteristics. While multiple distributions can be considered, as shown in the application of control of multiple agent distributions, completely heterogeneous systems may not be easily treated within the current DOC framework. Furthermore, for the DOC approach to be applicable in its present form, the agent performance must be expressed as a functional of the agent PDF. For systems that obey these simplifying assumptions, the DOC approach significantly reduces computational cost compared to classical optimal control methods. As a result, DOC is found to be computationally feasible for large numbers of agents without decoupling agent behaviors or averaging agent dynamics and, thus, without sacrificing optimality or completeness.

ACKNOWLEDGMENT

This research was funded by the ONR Code 321.

AUTHOR INFORMATION

Silvia Ferrari (ferrari@cornell.edu) is a professor of mechanical and aerospace engineering at Cornell University. Prior to that, she was professor of engineering and computer science, and founder and director of the NSF Integrative Graduate Education and Research Traineeship and Fellowship program on Wireless Intelligent Sensor Networks at Duke University. She is the director of the Laboratory for Intelligent Systems and Controls, and her principal research interests include robust adaptive control of aircraft, learning and approximate dynamic programming, and optimal control of mobile sensor networks. She received the B.S. degree from Embry-Riddle Aeronautical University and the M.A. and Ph.D. degrees from Princeton University. She is a Senior Member of the IEEE and a member of ASME, SPIE, and AIAA. She is the recipient of the ONR Young Investigator award (2004), the NSF CAREER award (2005), and the Presidential Early Career Award for Scientists and Engineers (PECASE) award (2006). She can be contacted at 214 Upton Hall, Cornell University, Ithaca, NY, 14853 USA.

Greg Foderaro received the B.S. degree in mechanical engineering from Clemson University in 2009 and the Ph.D. degree in mechanical engineering and materials science from Duke University in 2013. He is currently a staff engineer at Applied Research Associates, Inc. His research interests are in underwater sensor networks, robot path planning, multiscale dynamical systems, pursuit-evasion games, and spiking neural networks. He is a Member of the IEEE.

Pingping Zhu received the B.S. degree in electronics and information engineering and the M.S. degree in the institute for pattern recognition and artificial intelligence from the Huazhong University of Science and Technology and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Florida. He is currently a research associate in the Department of Mechanical and Aerospace Engineering at Cornell University. Prior to that, he was a postdoctoral associate at the Department Of Mechanical Engineering and Material Science at Duke University. His research interests include approximate dynamic programming, optimal control of mobile sensor networks, signal processing, machine learning, and neural networks. He is a Member of the IEEE.

Thomas A. Wettergren received the B.S. degree in electrical engineering and the Ph.D. degree in applied mathematics, both from Rensselaer Polytechnic Institute. He joined the Naval Undersea Warfare Center in Newport in 1995, where he has served as a research scientist in the torpedo systems, sonar systems, and undersea combat systems departments. He currently serves as the U.S. Navy senior

technologist for operational and information science, with a concurrent title as a senior research scientist at the center. He also is an adjunct professor of mechanical engineering at Pennsylvania State University. His personal research interests are in planning and control of distributed systems, applied optimization, multiagent systems, and search theory. He is a Senior Member of the IEEE and a member of the Society for Industrial and Applied Mathematics.

REFERENCES

- [1] J. Sanchirico and J. Wilen, "Optimal spatial management of renewable resources: Matching policy scope to ecosystem scale," *J. Environ. Econ. Manag.*, vol. 50, no. 1, pp. 23–46, 2005.
- [2] J. H. Reif and H. Wang, "Social potential fields: A distributed behavioral control for autonomous robots," *Robot. Auton. Syst.*, vol. 27, no. 3, pp. 171–194, 1999.
- [3] T. Balch and R. C. Arkin, "Behavior-based formation control for multi-robot teams," *IEEE Trans. Robot. Automat.*, vol. 14, no. 6, pp. 926–939, 1998.
- [4] J. P. Desai, J. Ostrowski, and V. Kumar, "Modeling and control of formations of nonholonomic mobile robots," *IEEE Trans. Robot. Automat.*, vol. 17, no. 6, pp. 905–908, 2001.
- [5] V. Gazi and K. M. Passino, "Stability analysis of social foraging swarms," *IEEE Trans. Syst., Man, Cybern.*, vol. 34, no. 1, pp. 539–557, 2004.
- [6] L. P. F. Giulietti and M. Innocenti, "Autonomous formation flight," *IEEE Contr. Syst. Mag.*, vol. 20, no. 6, pp. 3444–3457, 2000.
- [7] N. E. Leonard and E. Fiorelli, "Virtual leaders, artificial potentials and coordinated control of groups," in *Proc. Conf. Decision Control*, 2001, pp. 2968–2973.
- [8] A. Singh and J. P. Hespanha, "Moment closure techniques for stochastic models in population biology," in *Proc. IEEE American Control Conf.*, 2006, pp. 4730–4735.
- [9] A. Singh and J. P. Hespanha, "A derivative matching approach to moment closure for the stochastic logistic model," *Bulletin Math. Biol.*, vol. 69, no. 6, pp. 1909–1925, 2007.
- [10] S. Peng and Z. Wu, "Fully coupled forward-backward stochastic differential equations and applications to optimal control," *SIAM J. Contr. Optim.*, vol. 37, no. 3, p. 825–843, 1999.
- [11] S. T. Li and J.-F. Zhang, "Asymptotically optimal decentralized control for large population stochastic multiagent systems," *IEEE Trans. Automat. Contr.*, vol. 53, no. 7, pp. 1643–1660, 2008.
- [12] S. Thrun, M. Bennett, and W. Burgard, "Finding and optimizing solvable priority schemes for decoupled path planning techniques for teams of mobile robots," *Robot. Auton. Syst.*, vol. 41, no. 2, pp. 89–99, 2002.
- [13] S. M. LaValle and S. A. Hutchinson, "Optimal motion planning for multiple robots having independent goals," *IEEE Trans. Robot. Automat.*, vol. 14, no. 6, pp. 912–925, 1998.
- [14] C. C. Cheah, S. P. Hou, and J. J. E. Slotine, "Region-based shaped control for a swarm of robots," *Automatica*, vol. 45, no. 10, pp. 2406–2411, 2009.
- [15] S. D. Bopardikar, F. Bullo, and J. P. Hespanha, "A cooperative homicidal chauffeur game," *Automatica*, vol. 45, pp. 1771–1777, July 2009.
- [16] M. Huang, P. E. Caines, and R. P. Malhame, "The NCE (mean field) principle with locality dependent cost interactions," *IEEE Trans. Automat. Contr.*, vol. 55, no. 12, pp. 2799–2805, 2010.
- [17] 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," *Commun. Inform. Syst.*, vol. 6, no. 3, pp. 221–252, 2006.
- [18] G. Foderaro, S. Ferrari, and T. A. Wettergren, "Distributed optimal control of sensor networks for dynamic track coverage," submitted for publication.
- [19] G. Foderaro, S. Ferrari, and T. Wettergren, "Distributed optimal control for multi-agent trajectory optimization," *Automatica*, vol. 50, no. 1, pp. 149–154, 2013.
- [20] K. Rudd, G. Foderaro, and S. Ferrari, "A generalized reduced gradient method for the optimal control of mobile robotic networks," submitted for publication.
- [21] R. F. Stengel, *Optimal Control and Estimation*. New York: Dover Publications, Inc., 1986.
- [22] M. Athans and P. Falb, *Optimal Control*. New York: McGraw-Hill, 1966.
- [23] J. Betts, "Survey of numerical methods for trajectory optimization," *J. Guidance, Contr., Dyn.*, vol. 21, no. 2, pp. 193–207, 1998.
- [24] P. Gill, W. Murray, and M. Wright, *Practical Optimization*. Cambridge, MA: Academic, 1981.
- [25] (2004). Mathworks, Matlab Optimization Toolbox. [Online]. Available: <http://www.mathworks.com>
- [26] G. Huntington and A. Rao, "Optimal reconfiguration of spacecraft formations using a Gauss pseudospectral method," *J. Guidance, Contr., Dyn.*, vol. 31, no. 3, pp. 689–698, 2007.
- [27] P. Williams, "A comparison of differentiation and integration based direct transcription methods," in *Proc. AAS/AIAA Spaceflight Mechanics Meeting*, Copper Mountain, CO, Jan. 2005.
- [28] C. R. Hargraves and S. W. Paris, "Direct trajectory optimization using nonlinear programming and collocation," *J. Guidance, Contr. Dyn.*, vol. 10, no. 4, pp. 338–342, 1987.
- [29] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications* (Graduate Studies in Mathematics, vol. 112). Providence, RI: Amer. Math. Soc., 2010.
- [30] I. Kevrekidis, C. Gear, J. Hyman, P. Kevrekidis, O. Runborg, and C. Theodoropoulos, "Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis," *Commun. Math. Sci.*, vol. 1, no. 4, pp. 715–762, 2003.
- [31] J. P. Boyd, *Chebyshev and Fourier Spectral Methods, II Ed.* New York: Dover, 2001.
- [32] X. Mao, *Stochastic Differential Equations and Applications*. Chichester, U.K.: Horwood Publishing, 1997.
- [33] C. Fox, *An Introduction to the Calculus of Variations*. New York: Dover Publications, Inc., 1987.
- [34] L. Biegler, *Large-Scale PDE-Constrained Optimization*, (series Lecture Notes in Computational Science and Engineering, no. 30). Berlin, Germany: Springer-Verlag, 2003.
- [35] K. Rudd and S. Ferrari, "A constrained integration (CINT) approach to solving partial differential equations using artificial neural networks," *Neurocomputing*, vol. 155, pp. 277–285, May 2015.
- [36] M. Kaazempur-Mofrad and C. Ethier, "An efficient characteristic Galerkin scheme for the advection equation in 3-D," *Comput. Methods Appl. Mech. Eng.*, vol. 191, no. 46, pp. 5345–5363, 2002.
- [37] T. K. Sengupta, S. B. Talla, and S. C. Pradhan, "Galerkin finite element methods for wave problems," *Sadhana*, vol. 30, no. 5, pp. 611–623, 2005.
- [38] J. C. Tannehill, D. A. Anderson, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*. New York: Taylor and Francis, 1997.
- [39] J. C. McLachlan, *Finite Mixture Models*. Hoboken, NJ: Wiley, 2000.
- [40] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 2007.
- [41] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. Hoboken, NJ: Wiley, 1992.
- [42] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. Hoboken, NJ: Wiley, 1992.
- [43] Y. Hu, J.-G. Lou, H. Chen, and J. Li, "Distributed density estimation using non-parametric statistics," in *Proc. Int. Conf. Distributed Computing Systems*, June 2007, p. 28.
- [44] O. B. M. Jelasity and A. Montesor, "Gossip-based aggregation in large dynamic networks," *ACM Trans. Comput. Syst.*, vol. 23, no. 3, pp. 219–252, 2005.
- [45] J. C. Latombe, *Robot Motion Planning*. Norwell, MA: Kluwer, 1991.
- [46] C. Samson, "Time-varying feedback stabilization of car-like wheeled mobile robots," *Int. J. Robot. Res.*, vol. 12, no. 1, pp. 55–64, 1993.
- [47] A. Astolfi, "On the stabilization of nonholonomic systems," in *Proc. IEEE Conf. Decision Control*, 1994, pp. 3481–3486.
- [48] K. D. Do, Z.-P. Jiang, and J. Pan, "A global output-feedback controller for simultaneous tracking and stabilization of unicycle-type mobile robots," *IEEE Trans. Robot. Automat.*, vol. 20, no. 3, pp. 589–594, June 2004.
- [49] A. Bloch, M. Reyhanoglu, and N. McClamroch, "Control and stabilization of nonholonomic dynamic systems," *IEEE Trans. Automat. Contr.*, vol. 37, no. 11, pp. 1746–1757, Nov. 1992.
- [50] K. Pathak and S. Agrawal, "An integrated path-planning and control approach for nonholonomic unicycles using switched local potentials," *IEEE Trans. Robot.*, vol. 21, no. 6, pp. 1201–1208, 2005.
- [51] W. Lu, G. Zhang, and S. Ferrari, "An information potential approach to integrated sensor path planning and control," *IEEE Trans. Robot.*, vol. 30, no. 4, pp. 919–934, 2014.
- [52] J. E. Hopcroft, J. T. Schwartz, and M. Sharir, "On the complexity of motion planning for multiple independent objects; PSPACE-hardness of the 'warehouseman's problem,'" *Int. J. Robot. Res.*, vol. 3, no. 4, pp. 76–88, 1984.
- [53] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ: Wiley, 1991.

