Journal Title
XX(X):1–14
©The Author(s) 2016
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Multi-agent Policy Optimization: Optimality, Robustness, Safety, and Generalized Nash Equilibrium.

Etiosa Omeike[1], Jaime Fisac[1], Radhika Nagpal[1], and Lekan Molu[2]

Abstract

We consider continuous-time, nonlinear multi-agent trajectory optimization under state and input constraints. With our eyes toward optimality, robustness, and safety of the ensuing policy we combine game-theoretic, control-theoretic and learning-based approaches in a generalized Nash Equilibrium framework. Here, global trajectories are a consequence of the collective emergent behavior from individual agents. All agents possess local sensory information only; this informs self-organization into disjoint subgroups – the union of these separate disjoint structures constitute the whole group behavior. Control laws for topological subgroup formation and group geometric kinematics are computed from Hamilton-Jacobi equations via parallel optimization on graphic processing units. We formulate the global group heading and geometric cohesion behavior problem as separable reachability and distributed trajectory optimization problems within a *generalized Nash Equilibrium convergence* objective. One has the intuitive feeling that this safe multiagent trajectory optimization with Nash equilibrium seeking global objective may inspire safe robust autonomy among a complex network of aerial, ground, or underwater multiagent systems in the real-world in the future.

Keywords
Reachability analysis; multiagent systems; multiphase flows; levelset methods.

Introduction

This paper is devoted to several themes, mostly concerning numerical algorithms for determining *globally safe and optimal trajectories that are Nash-equilibrium seeking* in networks of complex multiagent systems characterized by state and input constraints. We take inspiration from natural swarms, particularly the murmuration of European starlings (*sturnus vulgaris*). In these settings, local flocks within large murmurations maintain an anisotropic formation based on a topological interaction, regardless of sparsity of birds on a phase space (Cavagna et al. 2010). Thus, intra- and inter-flock collisions are avoided and attacks are fended off (Ballerini et al. 2008). Approximating the viscosity solutions of nonconvex Hamilton-Jacobi partial differential equations (that we use to characterize these vector fields) with proximal operators and importance sampling, the Hamiltonian, control laws, and strategies that govern the transient behaviors of many systems that possess structural subsystems with unique nearest neighbor properties are computed.

Through empirical (Ballerini et al. 2008; Cavagna et al. 2010; Helbing et al. 2000; Vicsek et al. 1995; Bialek et al. 2012) and theoretical findings (Jadbabaie et al. 2003), evidence now abounds that in certain natural species that exhibit collective behavior, convergence and group cohesion is based on simple topological interaction rules that they employ to keep a tab on one another in *local flocks* for collision avoidance, preserving density and structure in an anisotropic formation, and exhibiting flock splitting, vacuole, cordon, and flash expansion isotropically (Haiken 2021). This aids these animals in emerging an eye-pleasing local anisotropic synchrony, which taken together among possibly hundreds of thousands of local interactions[*] (Haiken 2021), keep these animals whirling, swooping, and flying in isotropic formations (Ballerini et al. 2008). Thus, individual agents aggregate into substructures within the overall system, and overall group motion is synergized via local topological interactions so that a stable global heading and cohesion (Jadbabaie et al. 2003) is preserved.

There exists evidence that when an individual within a subgroup of starlings senses danger (e.g. an attack from a Peregrine Falcon), it changes its course immediately. Owing to the lateral vision in such animals, immediate *nearest neighbors* change course in response. This information is propagated across the entire group within the fraction of a second (Ballerini et al. 2008), resulting in elegant formations (Haiken 2021).

In this paper, borrowing inspiration from these natural behaviors, we will introduce a reachable safety framework for distributed trajectory optimization with optimality guarantees. Our lead stems from the fact that the interaction among various subgroups into which individual agents self-organize (Ballerini et al. 2008; Cavagna et al. 2010) and overall motion can be characterized by dynamically evolving interfaces that separate the respective multiple regions or phases (Kim et al. 2010; Zaitzeff et al. 2019;

[1]Princeton University [2]Microsoft Research.

Corresponding author:
Lekan Molu, Microsoft Research NYC, Microsoft Corporation, 300 Lafayette Street, New York City, New York, 10012, USA.
Email: lekanmolu@microsoft.com

[*] It has been reported that no birds fly together with greater coordination and complexity than European starlings, with murmurations counting upwards of 750,000 individual birds!

Karnakov et al. 2021). These systems exhibit separable structural dynamics that altogether dynamically evolve in a manner akin to the murmuration of *European starlings, fish schools in the ocean, multiphase flows, or crashing ocean waves at mesoscale*. For large-scale multiagent (robotic) autonomous systems with characteristic topological changes among subgroup boundaries and possessing structural intra-group cohesion e.g. (Saravanos et al. 2022), we set out to replicate some of these natural behavior and improve state of-the-art optimization algorithms in the autonomous systems that we design and build by managing state and input constraints in the nonlinear regime with notions of guaranteed safety during design and synthesis of controllers. We leverage levelset theory (Sethian 2000) on natural collective behaviors, and extend the reachability analysis of (Mitchell et al. 2005) for safe, distributed, and large-scale multiagent systems with separable structural kinematics and dynamics.

The body of this paper is structured as follows: we present background and motivation for our work in § and formally introduce our safe trajectory optimization schemes in § .

Preliminaries

In this section, we review related works on large-scale optimization of a team of multi-agent systems, provide a motivation for our approach, and then introduce a list of common notations used throughout the rest of the article.

Motivation

Throughout this article, *safe trajectories* or *safety* shall mean computed trajectories, particularly from *Hamilton-Jacobi representations*, that satisfy *all* input and state constraints *within a time interval*. Formally by safety, we mean thre freedom of a system from harm (Hobbs et al. 2023). We loosely use the terms *decision* to mean *control law* or *controller*. The *utility function* that quantifies the state in which an agent exists may be referred to as a cost, objective, or reward function. At issue is the *safe* trajectory optimization for large-scale multiagent systems with aggressive paths-following requirement towards eventual collective equilibrium of global state. This requirement necessitates a principled management of constraints so as to avoid hazard during real-world deployment.

This may be considered a constrained optimization problem involving minimizing the stage and terminal costs in a Hamilton-Jacobi function setting subject to safety constraints e.g. in augmented Lagrangian settings where inter-agents separation, optimality of local control laws, preservation of global formation are a requirement. A simplistic distributed optimization modeling concept, does not lend itself effectively to global equilibrium of collective states in a principled variational, operator-theoretic, or control-theoretic methods. A principled way to consider the solution concept would be to model the decision-making problem as games played between interacting agents where *strategic interactions* occur over networks and where collective states must arrive at a *(generalized) Nash Equilibrium* (GNE)– essentially an equilibrium state for the whole system where each agent's control law is optimal *with respect to other agent's controllers* and no single agent gains an advantage by unilaterally amending its own decision.

Here, centralization and decentralization of communication must be carefully managed. For centralized communication, the information structure implies that all agents update their state simultaneously. Over wireless sensor networks, however, challenges such as limited communication bandwidths in occluded operational environments imposes a hard constraint. A semi- or wholly-decentralized communication strategy with, for example, local neighbor rules, makes principled multiagent control and optimization feasible in large-scale systems that requires no global information over the entire network. Ideally, the goal is to plan with collective cooperation; however, this imposes stringent communication throughput requirements. We thus drop the requirement that an external agent enforce cooperation so that a *noncooperative game ensues*. The challenge, as opposed to distributed large-scale trajectory optimization e.g. (Saravanos et al. 2022), is that the separability of collective objective function is not assumed as such formulations at best only enforce constraints softly via the Karush-Kuhn-Tucker (KKT) conditions. These do not necessarily provide guaranteed safety constraints satisfaction *in the large*.

In our formulation, each agent has independent control over its own objective, subject to the the *control law* of its neighboring agents. Hence, the notion of *best response* in coupled multi-agent decision settings where partial authority applies. And we adopt a *dynamic game model* so that the game evolves in time as players implement their policies during play – as this is well-known to cover decision-making problems that are the bane of emerging inter- and multi-disciplinary problems (Li et al. 2022).

The applications of the formalisms we introduce herein extend beyond murmurations. This can be used to model the interactive behavior of independent decision-making agents on social network services, network of distributed autonomous vehicles, search-and-rescue aerial robots, consensus formation of connected systems, bioinspired complex systems that model the complex movement of muscular hydrostats, smart grids' independent station generators, and the internet of things among others.

Notations

Throughout, capital and lower-case Roman letters respectively denote matrices and vectors while calligraphic letters denote sets. Exceptions: time variables e.g. t, t_0, t_f, T will always be real numbers. Unless otherwise stated, vectors $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t)$ are reserved for admissible control (resp. disturbance) at time t. We say $\boldsymbol{u}(t)$ (resp. $\boldsymbol{v}(t)$) is piecewise continuous in t, if for each t, $\boldsymbol{u} \in \mathcal{U}$ (resp. $\boldsymbol{v} \in \mathcal{V}$), $\mathcal{U}($ resp. $\mathcal{V})$ is a Lebesgue measurable and compact set. At all times, any of u or v will be under the influence of a *player* such that the motion of a state x will be influenced by the coercion of that player. Our theater of operations is that of conflicting objectives between players – so that the problem at hand assumes that of a *game*. And by a game, we do not necessarily refer to a single game, but rather a collection of games. Each player in a game will constitute either a pursuer (P) or an evader (E).

We let $\{\lambda_i(X)\}_{i=1}^n$ denote the n-eigenvalues of $X \in \mathbb{R}^{n\times n}$ with $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. When an optimized variable, u, is optimal with respect to an index of performance, it shall be denoted $u^\star$. All vectors are column-stacked. The norm $\|X\|$ of a matrix X is sup $\|X\|$ over $\|X\| = 1$. We define Ω as the open set in $\mathbb{R}^n$. To avoid the cumbersome phrase "the state x at time t", we associate the pair (x, t) with the system's *phase*. The Cartesian product of Ω and the space $T = \mathbb{R}^1$ of all time values is called the *phase space*. The interior of Ω is denoted by $\text{int}\ \Omega$; whilst the closure of Ω is $\bar{\Omega}$. Let $\delta\Omega\ (:= \bar{\Omega}\backslash \text{int}\ \Omega)$ be the boundary of Ω.

Definitions

The set of players in a game is denoted $\mathcal{N} = \{i, j, \ldots\}$ with subscript indexing indicating players e.g. $\mathcal{N}_i$ for player i. The set of neighbors of player i is $\mathcal{N}(i) \subseteq \mathcal{N}$. Player i moves dynamically with a control $u_i \in \pi_i$ (i.e. u_i belongs to a policy class π_i) that is both (a) optimal with respect to its own objective $\mathcal{J}_i$; and (b) optimal with respect to its neighboring players' current policy $\pi_{-i} \in \Pi_{j \in N, j \neq i}$.

Trajectory Optimization

To mimic these on real-world systems, autonomous agents must operate near actuator and sensing limits under demanding control computations. A common direction is to train a neural perceptive controller offline to follow a set of predetermined waypoints which is then deployed online (Song et al. 2021; Torrente et al. 2021). Other approaches generate optimal trajectories offline under different initial conditions for the nonlinear problem from which a feedback controller in a Gaussian process regression setting is recovered (Gerdts et al. 2009). Simulation of optimal policies with deep reinforcement learning (see for example, (Vinitsky et al. 2022)) for motion-planning, or trajectory optimization is another popular approach.

Static optimization is a popular approach for solving these problems. Here, the choice of a sampling distribution for the decision variables must be chosen by a user ahead of the optimization process. Typical choices include thermodynamics-inspired methods such as simulated annealing (Kirkpatrick et al. 1983), evolutionary-based genetic algorithms (Holland 1975), or cross-entropy probabilistic sampling-based planning methods (Kobilarov 2012) *inter alia*. However, for nonlinear problems with often non-convex costs, maintaining the consistent decrement in value function at every iterate (often) subject to input and state constraints becomes challenging as problem size grows. When there are non-smooth constraints in the state space, the problem is NP-complete (LaValle 2006). The sampling distribution associated with these methods is critical for regions' exploration, coverage, and the importance of each sample to the problem solution is often chosen on either heuristics or meta-heuristics. However, these sampling-based methods are computationally intractable for real-time high-dimensional systems applicable to modern distributed agent optimization and control. Given the nonlinearity associated with most multi-agent robotic problems, gradient-based optimization introduces multiple local minima; and the presence of obstacles in the state space may hinder convergence guarantees since special differentiation techniques may have to be adopted in such instances (Clarke et al. 2008).

Adjacent to dynamic optimization schemes are sampling-based algorithms. Here, the problem is repeatedly solved for gathered samples from rolled-out trajectories over a time period. Path integral control (Kappen 2005) and its variants (Theodorou et al. 2010; Williams et al. 2016) play a notable role in stochastic sampling-based optimization algorithms. Employing the stochastic version of the Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE) for affine control systems with an additive Brownian motion, the PDE is linearized into a form of the backward Chapman-Kolmogorov equation (Pavliotis 2014) after an exponential transformation of the value function. An application of the Feynman-Kac Lemma (Pavliotis 2014; Leonid Koralov 2007) naturally then follows, so that the linearized PDE admits *path-integral* representations of the value function whose derivatives yield the optimal controller. These schemes possess more robustness whilst being less prone to local minima compared to gradient-based algorithms. These methods are attractive because they are derivative-free when optimizing the cost with respect to the dynamics (Williams et al. 2017). A key drawback of sampling-based optimization techniques, however, is their lack of scalability to large multiagent systems — with the increase in number of agents comes the difficulty of sufficient exploration in the state space based on sampled trajectories.

The foregoing methods, while yielding impressive empirical results, lack safety guarantees. For example, deep RL methods are unable to combine the long-term spatial reasoning of neural network approximators with the precision requirement of fine motor control, the dynamic manipulation, and the sturdy dexterity requirement of real-world systems. Sampling-based planners often employ centralized optimization methods where a centralized controller carries out all computation needed for actuation. In both cases, needed controllers typically exist at multiple timescales that these controllers do not account for at execution time. This lack of multiscale, hierarchical information processing, extraction, and alignment of temporally-structured reasoning for life-long control, adaptation, and memory shuffling has necessitated the timely study of safe optimal trajectories for large-scale systems.

Closest in spirit to the algorithms we present here are the nested-distributed (ND-DDP) and merge-distributed (MD-DDP) differential dynamic programming algorithms of (Saravanos et al. 2022), whereupon utilizing augmented Lagrangian forms in managing state and input constraints among multiple agents with neighborhood rules, the decentralization – informed by local communication among agents (provided by the introduced parallelization scheme with the alternating direction method of multipliers algorithm) – enabled control optimization of up to $4,096$ cars with $16,384$ states with the MD-DDP version of the algorithm. It should be noted, however, that these algorithms merely utilize primal feasibility (for constraint satisfaction) and dual feasibility (for optimality guarantees). At best, these constraints are soft in nature (owing to ADMM distribution scheme) so that a safe solution for the overall system is not guaranteed even if local controllers enforce local safety constraints (this is the case even for controllers that enforce strict constraints such as sequential quadratic programming methods). While empirical checks can be made for the primal and dual residuals of the ADMM solver to test for safety, this is not inculcated into the algorithm to demonstrate safe controller optimization throughout the optimization lifecycle.

In addition, this algorithm offers significant computational complexities beyond the vanilla quadratic computation requirements of traditional DDP algorithms – arising out of the matrix inversion requirements of DDP and its scaling to the overall number of agents (Saravanos et al. 2022).

Representation of Hamilton-Jacobi Equations

A principled way for finding globally optimal controllers for a trajectory optimization problem is to consider Cauchy-type Hamilton-Jacobi equations of the evolution form

$$\phi_t + H(x, t, \phi, D\phi) = 0, \tag{1}$$

where $D\phi$ is the spatial gradient of the vector field ϕ. As the solution to these equations generally lack suitable existence and uniqueness properties, a numerically means of finding feasible solutions is to adopt the numerical levelset representation of the solution (Crandall and Lions 1984a; Adalsteinsson and Sethian 1999; Sethian 2000) which provide numerically consistent and monotone solutions to HJ problems (Osher and Shu 1987). Suppose that we are given multiple regions $\mathcal{S}_i, \ldots, \mathcal{S}_n$, each separated by interfaces Γ_i (i.e. Γ_i separates region $\mathcal{S}_i$ from $\mathcal{S}_{i+1}$), and where each interface Γ_i moves with a given speed V_i in its normal direction. For n regions, the collective interface is an $n-1$ dimensional hypersurface in $\mathbb{R}^n$. Embedding this interface as the zero-level set of an unsigned distance function $\phi(x)$ to the interface Γ_i, where $\phi(x)$ is zero on the interface and positive everywhere else to the boundary. We are concerned with a numerical algorithm for evolving $\phi(x)$.

While these methods have attractive numerical robustness, convergence, and accuracy characteristics (Mitchell et al. 2005; Fisac et al. 2019a,b; Herbert et al. 2017), they do not scale to higher dimensions owing to the intrinsic space-time discretization that grows exponentially in size as problem dimensions increase. *What if the problem admits special structure (such as nearest-neighbor behaviors)* that we are able to leverage ideas from computing flows of multiphase regions separated by interfaces (Karnakov et al. 2021) or the evolution of mean curvatures (Osher and Fedkiw 2004), fronts of multiphase fluids (Osher and Sethian 1988), or advection of substrates in multiple region fluids (Saye and Sethian 2011; Sethian 1996; Zaitzeff et al. 2019) in computing optimal trajectories? This is the central theme of this paper.

The key idea is that when the multi-agent problem has separable structure, one can resolve the associated HJ equations in a numerically consistent manner by considering the interface of these separable structures, i, as evolving dynamic interfaces, Γ_i, via the modified levelset equation (1)

$$\phi_t + \Gamma_i H(x, t, \phi, D\phi) = 0 \tag{2}$$

At each time step, we advance each interface ϕ_i for agents in a region i by solving the levelset equation for a small time step Δt; afterwards, we reconstruct the unsigned distance function that captures the safe set, which informs the controller optimization for the agents $\mathcal{S}_i \in \mathcal{C}$. Note that contrary to distributed consensus-type algorithms such as (Saravanos et al. 2022), this framework guarantees safe exploration by directly incubating safety into the trajectory optimization problem.

Using the terminal function in Hamilton-Jacobi-Isaacs functionals and backward reachability theory, we construct a theorem for computing "safety-preserving" BRATs (Mitchell 2001; Mitchell et al. 2005; Mitchell 2020) across local regions. BRATs are those zero-level sets (Sethian 2000) of implicitly-defined value functions on a state space that return a "safety-satisfying" certificate after solving a time-dependent Hamilton-Jacobi Isaacs equation (Evans and Souganidis 1984; Crandall and Lions 1983). TO-DO: We then aggregate the over-approximated BRATs of the respective subsystems.

Differential optimal control theory and games provide a useful paradigm for the verification of multiple agents interacting over a vectograme.g. $\mathbb{R}^n$ (Isaacs 1965)'s verification theorem. Both rely on the resolution of the Hamilton-Jacobi-Bellman (HJB) equation or its Isaacs counterpart (HJI). As HJ-type equations are seldom regular enough to admit a classical solution for almost all *practical* problems, "weaker" or "viscosity" solutions (Lions 1982; Evans and Souganidis 1984; Crandall et al. 1984) provide generalized solutions to HJ partial differential equations (P.D.E.s) under relaxed regularity conditions; these viscosity solutions are not necessarily differentiable anywhere in the state space, and the only regularity prerequisite in the definition is continuity (Crandall and Lions 1983). However, wherever they are differentiable, they satisfy the values of HJ P.D.E.s in a classical sense. Thus, they lend themselves well to many real-world problems existing at the interface of discrete, continuous, and hybrid systems (Evans and Souganidis 1984; Lygeros 2004; Mitchell et al. 2005; Mitchell 2020).

With the elegant theoretical results of (Crandall and Lions 1984b,a; Evans and Souganidis 1984; Osher and Shu 1991; Crandall et al. 1984), stable essentially non-oscillatory Lax-Friedrichs numerical integration schemes provide *consistent* and *monotone* viscosity solutions with high accuracy and precision *on a mesh* to multi-dimensional HJ-type equations (Consistent solutions to HJ equations are those whose explicit marching schemes via discrete approximations to the HJ IVP agree with the nonlinear HJ solution (Crandall and Lions 1984a). Such schemes are said to be *monotone* e.g. on $[-\mathbb{R}, \mathbb{R}]$ if the numerical approximation to the vector field of interest is a non-decreasing function of each argument of the discrete approximation to the vector field.) By consistency, we mean that the numerical approximation to the HJ equation agrees with a defined HJ initial value problem; and by monotonicity, we mean the explicit marching schemes an Hamiltonian's flux are a non-decreasing function (for a 1-D case) of each argument of the vector field upon which the system is defined. For more details, see (Crandall and Lions 1984a; Osher and Shu 1991). However, by discretizing the Hamiltonian on a dimension-by-dimension basis, the scheme suffers from scalability as a result of the exponential computational complexity associated with grid resolutions of value functions (Herbert et al. 2021; Bajcsy et al. 2019; Bansal and Tomlin). Given the limits of computational resources and memory when resolving practical problems for multi-agent systems, what if we exploit local structures within a complex system and resolve the overall Hamiltonian by an aggregation of the computation of the numerical fluxes of local Hamiltonians? This is the central question that this paper seeks to address.

The motion of interfaces between subgroups are cast as zero level set of implicitly defined unsigned distance functions UDFs. Interfaces evolve by solving time-dependent Eulerian initial value partial differential equations of these SDFs in the form of viscosity solutions (Crandall and Lions 1983) to hyperbolic conservation laws (Crandall and Majda 1980), which are essentially the original HJ equations (Evans and Souganidis 1984). Hence, the front's position gets updated by this means and the interface velocity is derived from physics on and off the interface.

Popular numerical tools for simulating the evolution of interfaces in applied physics, mathematics, and computational sciences include fast marching front tracking methods (Sethian 1987, 1996), Voronoi implicit interface methods (or VIIMs) (Saye and Sethian 2011; Zaitzeff et al. 2019), discretization schemes for Hamilton-Jacobi equations (Tsitsiklis 1995), or multilayer volume-of-fluid methods (Karnakov et al. 2021) for foaming across scales. While each of these schemes has its own advantages, we will resort to levelset methods (Sethian 2000) in developing algorithmic efficiency for the trajectory optimization among a *collection* ($\mathcal{C}$) of *disjoint* multiple agents *sets*, $\mathcal{S} \subseteq \mathcal{C}$. The entire collection of agents moves over an open set $\Omega \subset \mathbb{R}^n$. When subgroups $\{\mathcal{S}_i\}_{i=1}^{i=k} \subset \mathcal{C}, k < n$ of aerial agents must traverse a narrow opening, or temporarily break apart to avoid collisions, for example, we expect topological changes to occur without explicit surgery. Furthermore, group geometric formation and cohesion are achieved by local changes to subgroups' collective heading and speed. This is a challenging problem numerically and we look to natural behaviors in animals and multiphase simulations for inspiration. In our problem construction, a *finite set* of individual agents self-organize into local structural groups (which we refer to as subgroups, $\mathcal{S}$); subgroups interact based on nearest neighbor rules so that effective group consensus is dictated by interactions among separate subgroups.

Throughout, our theater of operations involve multiple aerial Dubins vehicles (Dubins 1957), treated as kinematic models that possess linear and angular speeds as state variables. The trajectory optimization problem is played as a game of multiple vehicles in a three-dimensional space. As we are dealing with multiple agents, objectives such as collision avoidance, agents' spatial separation, overall group coherence become paramount. At issue is continually computing the points set that belong on the *reachable set* boundary as the game advances forward in time. These reachable sets are those state space subsets where agents may collide into one another, hit obstacles, lose topological integrity or miss group coherence behavior. We

As all agents move in $\Omega \subset \mathbb{R}^n$, we want a stable and safe numerical algorithm that correctly represents dynamic interface boundary conditions, accurately represents kinematics, whilst sensitive enough to rapidly interpret subgroup topological and structural changes. This *safe* trajectory optimization problem has been considered previously in several communities. In the control community, (Tsitsiklis 1995) worked under the restrictive assumption that HJ's running cost is independent of the control laws' update – providing an $O(n/p)$ parallel algorithm that resolved the discretized HJ-Hamiltonian for p processors on n grid points provided that $p = O(\sqrt{n}/\log n)$. In front tracking methods, a Lagrangian geometric representation techniques track the surfaces between separate structures with mechanisms such as triple point junctions which are endowed with shared nodes so that members of the front set are updated as time evolves (Kim et al. 2010). In applied mechanics and microfluidics, methods such as volume-of-fluid (Balcázar et al. 2015) resolve the interaction among multiple phases that are separated by thin boundaries with volume fraction fields for each region or unique functions in levelset methods – leading to coalescence prevention among multiple regions. However, they come at a computational cost of $\mathcal{O}(N_{regions}N_{cells})$. To improve its scalability, (Karnakov et al. 2021) compactly stored many fields thereby keeping needed scalar fields constant and independent of regions to simulate.

A few mathematical notations, conventions, and taxonomy used throughout the paper are in order at this juncture. Capital and lower-case Roman letters are matrices and vectors respectively. Exceptions: time variables such as t, t_0, t_f, T are real numbers throughout. Calligraphic letters are sets. Exception: the HJ equation's solution (shortly introduced) is the set $\Omega \in \mathbb{R}^n$ in which agents move. We work in a multi-agent system context where individual agents self-organize into phases or regions $\mathcal{S}$ which are in turn members of a union of multiple regions $\mathcal{C}$. Note that every $\mathcal{S} \subseteq \mathcal{C}$ and all members of $\mathcal{C}$ are disjoint from one another i.e. $\mathcal{S}_i \cup \mathcal{S}_j = \emptyset$ for any $i \neq j$. The total number of elements in $\mathcal{S}$ is denoted $[\mathcal{S}]$, and we denote by $\text{int}\,\Omega$ the interior of Ω. The closure of Ω is $\bar{\Omega}$. We let $\delta\Omega$ $(:= \bar{\Omega}\backslash \text{int}\,\Omega)$ be the boundary of Ω.

Structural homogeneity of agents' motion in every region $\mathcal{S}_i \in \mathcal{C}$ for $i = 1, \cdots, [\mathcal{C}]$ applies; this is enforced by introducing an external disturbance on the zeroth-index agent (this aids compactness of the zero levelset of an $\mathcal{S}_i$ as we will introduce shortly). Hence, each trajectory optimization episode can be characterized as a pursuit *game*, Γ. And by a game, we do not necessarily refer to a single game, but rather a *collection of games*, $\Upsilon = \{\Gamma_1, \cdots, \Gamma_g\}$. Such a game terminates when *capture* occurs, that is the distance between players falls below a predetermined threshold. Each player in a game shall constitute either a pursuer (P) or an evader (E). Let the cursory reader not interpret P or E as controlling a single agent. In our setup, we are poised with several pursuers (enemies) or evaders (peaceful citizens). However, when P or E governs the behavior of but one agent, these symbols will denote the agent itself. An evading agent in a region $\mathcal{S}_i$ has a state notation x_a^i (read: the state of agent a in region i). A state x_a^i has linear velocity components, $x_{a_1}^i, x_{a_2}^i$, and heading $x_{a_3}^i := w_a^i$. When we must distinguish an agent $x_a^i \in \mathcal{C}_x$ from some other agent e.g. in another multiphase $\mathcal{C}_y$, we shall write ${}^x x_a^i$ and ${}^y x_a^i$ respectively. Given the various possibilities of outcomes, the "best" outcome is resolved by a *payoff*, Φ, whose extremal over a time interval will constitute a *value*, V[†]. We adopt (Isaacs 1965)'s language so that if the payoff for a game is finite we shall have a *game of kind* (a qualitative game); and for a game with a continuum of payoffs we shall have *a game of degree* (quantitative games). The *strategy* executed by P or E during a game shall be denoted by $\alpha \in \mathcal{A}$ (resp. $\beta \in \mathcal{B}$).

[†] The functional $\mathbf{\Phi}$ may be considered a functional mapping from an infinite-dimensional space to the space of real numbers.

The many interacting subsystems under consideration employ (i) natural units of measurements that are the same for all agents; (ii) kinematics with same linear speeds but with a capacity for orientation changes; (iii) inter-region interaction occurs within unique and distinct state space manifolds; and by agents maneuvering their direction, a kinematic alignment is obtained with other regions; (iv) region-to-region interaction occurs when a pursuer is within a threshold of capturing any agent in a region; (v) the interaction among respective regions is described by the time-evolution of an interface, which is the zero-level set of the objective functional of the respective local subgroups.

Backward Reachability from Differential Games Optimal Control

We briefly review reachability theory and make a connection to the scope of this article. A basic characteristic of a control system is to determine the point sets within the state space that are *reachable* with a control input choice. An example objective in *reachability analysis* could be a target ($\mathcal{L}$) protection objective by an evading player E from a pursuing player P. Our treatment here is a special case of Isaac's homicidal chauffeur's game (Isaacs 1965), whereupon P and E travel at constant linear speeds but have different headings.

Backward reachability consists in avoiding an unsafe set of states under the worst-possible disturbance and at all times. In light of our multiple trajectory optimization goal, we want to find the set of reachable states that lie along the trajectories of the solution to a first order nonlinear HJI P.D.E. that originates from some initial state $x_0 = x(0)$ up to a specified time bound, $t = t_f$: from a set of initial and unsafe state sets, determine if there is an initial phase that the HJI PDE's solution enters an unsafe set. In general, we seek a terminal payoff $g(\cdot) : \mathbb{R}^n \to \mathbb{R}$ such that

$$|g(x(t))| \leq k, \quad |g(x(t)) - g(\hat{x(t)}) \,| \leq k|x(t) - \hat{x}(t)\,| \tag{3}$$

for constant k and all $T \leq t \leq 0$, $\hat{x}$, $x \in \mathbb{R}^n$, $u \in \mathcal{U}$ and $v \in \mathcal{V}$.

Suppose that a pursuer's mapping strategy (starting at t) is $\beta : \bar{\mathcal{U}}(t) \to \bar{\mathcal{V}}(t)$ provided for each $t \leq \tau \leq T$ and $\boldsymbol{u}(t), \hat{\boldsymbol{u}}(t) \in \bar{\mathcal{U}}(t)$; then $\boldsymbol{u}(\bar{t}) = \hat{\boldsymbol{u}}(\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$ implies $\beta[\boldsymbol{u}](\bar{t}) = \beta[\hat{\boldsymbol{u}}](\bar{t})$ a.e. on $t \leq \bar{t} \leq \tau$. Suppose further that the player P is controlling the strategy β and minimizing, while the player E is controlling its strategy, α, and maximizing. For any admissible control-disturbance pair $(\boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot))$ and initial phase (x_0, t_0), Crandall (Crandall and Lions 1983) and Evan's (Evans and Souganidis 1984) claim is that there exists a unique trajectory, $\boldsymbol{\xi}(t)$, the motion of the dynamical system, (5), passing through phase $(\boldsymbol{x}_0, t_0)$ under the action of control $\boldsymbol{u}$, and a worst-possible disturbance $\boldsymbol{v}$, and observed at a time t afterwards i.e.

$$\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t; t_0, x_0, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)). \tag{4}$$

Equation (4) is a solution of the following dynamical system, represented as a first-order p.d.e.

$$\dot{x}(\tau) = f(\tau, x(\tau), \boldsymbol{u}(\tau), \boldsymbol{v}(\tau)) \quad T \leq \tau \leq t, \quad x(t) = x, \tag{5}$$

almost everywhere (a.e.); where $f(\tau, \cdot, \cdot, \cdot)$ and $x(\cdot)$ are bounded and Lipschitz continuous. This bounded Lipschitz continuity property assures uniqueness of the system response $x(\cdot)$ to controls $\boldsymbol{u}(\cdot)$ and $\boldsymbol{v}(\cdot)$ (**?**). a.e. with the property that

$$\boldsymbol{\xi}(t_0) = \boldsymbol{\xi}(t_0; t_0, x_0, \boldsymbol{u}(\cdot), \boldsymbol{v}(\cdot)) = x_0. \tag{6}$$

In backward reachability analysis, the lower value of the differential game (**?**) is used in constructing an analysis of the backward reachable set (or tube). The differential game's lower value for a solution $x(t)$ that solves (5) for $\boldsymbol{u}(t)$ and $\boldsymbol{v}(t) = \beta[u](\cdot)$ is used in backward reachability analysis, given as

$$\begin{aligned} V^-(x,t) &= \inf_{\beta \in \mathcal{B}(t)} \sup_{\boldsymbol{u} \in \mathcal{U}(t)} \boldsymbol{\Phi}(\boldsymbol{u}, \beta[\boldsymbol{u}]) \\ &= \inf_{\beta \in \mathcal{B}(t)} \sup_{\boldsymbol{u} \in \mathcal{U}(t)} \int_t^T l(\tau, \boldsymbol{x}(\tau), \boldsymbol{u}(\tau), \beta[\boldsymbol{u}](\tau)) d\tau + g\left(\boldsymbol{x}(T)\right). \end{aligned} \tag{7}$$

Lemma 1. Theorem 1, (Mitchell et al. 2005). *The backward reachability problem resolves the infimum-supremum of the non-anticipative strategies of P and the controls of E as an extremum of the cost functional over a time interval (time of capture),* $t \in [-T, 0]$ *is given by*

$$\frac{\partial V^-}{\partial t}(\boldsymbol{x}, t) + \min\{0, H^-(t; \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, V_x^-)\} = 0, \; x \in \mathbb{R}^n, \tag{8a}$$

$$V^-(\boldsymbol{x}, 0) = g(\boldsymbol{x}), \tag{8b}$$

$$\text{where } H^-(t; x, \boldsymbol{u}, \boldsymbol{v}, p) = \max_{u \in \mathcal{U}} \min_{v \in \mathcal{V}} \langle f(t; x, \boldsymbol{u}, \boldsymbol{v}), p \rangle, \tag{8c}$$

and p, the co-state, is the spatial derivative of V^- *w.r.t x. where the vector field* V_x^- *is known in terms of the game's terminal conditions so that the overall game is akin to a two-point boundary-value problem.*

Henceforward, we will remove the negative superscript on the lower value and Hamiltonian (8).

Flock F_j and F_k within a murmuration, $F_j \cup F_k \cup F_l \cdots$ are separated by *partitions*, or *interfaces*, $\Gamma_{jk}, \Gamma_{kl}, \cdots$. This interface may be implicitly represented as a signed distance function $\Phi(x)$ which is negative on the interior of each flock, and zero on the edges. The zero-level set (i.e. $\Phi(x) = 0$) corresponds to the interface V (Sethian 1987). As the system evolves over time, F_j's interface Etiosa: why doesn't F_j have numerous interfaces? Interfaces are defined as the zero-level set of a signed distance function between two flocks, so I'd imagine that every flock has an interface with every other flock with a murmuration. (zero-level set) motion can be parameterized by time, so that the flow field $V(x,t)$ is equivalent to the solution of the Cauchy-type Hamilton Jacobi partial differential equation (Evans and Souganidis 1984; Crandall and Lions 1983):

$$V_t + v_j |\nabla V_j| = 0, \quad j = 1, \cdots, n_f, \tag{9}$$

where v_j is the flow speed for F_j. Equation (29) is the level set equation (Sethian 2000).

In the sentiment of (Mitchell et al. 2005), we say the zero sublevel set of $g(\cdot)$ in (8) i.e. $\mathcal{L}_0 = \{x \in \bar{\Omega} \,|\, g(x) \leq 0\}$, is the *target set* in the phase space $\Omega \times \mathbb{R}$ for a backward reachability problem (Mitchell 2001). This target set[‡] can represent the failure set, regions of danger, or obstacles to be avoided e.t.c. in the vectogram. And the *robustly controlled backward reachable tube* for $\tau \in [-T, 0]$[§] is the closure of the open set

$$\mathcal{L}([\tau,0],\mathcal{L}_0) = \{x \in \Omega \,|\, \exists\, \beta \in \bar{\mathcal{V}}(t)\, \forall\, \boldsymbol{u} \in \mathcal{U}(t), \exists\, \bar{t} \in [-T,0], \boldsymbol{\xi}(\bar{t}) \in \mathcal{L}_0\}, \, \bar{t} \in [-T,0]\,. \tag{10}$$

Read: the set of states from which the strategies β of P, and for all controls $\mathcal{U}(t)$ of E imply that we reach the target set within the interval $[-T, 0]$. More specifically, following Lemma 2 of (Mitchell et al. 2005), the states in the reachable set admit the following properties w.r.t the value function V

$$x \in \mathcal{L}_0 \implies V^-(x,t) \leq 0 \text{ and } V^-(x,t) \leq 0 \implies x \in \mathcal{L}_0. \tag{11}$$

Methods.

For the reader unfamiliar with reachability theory, a background necessary for processing the formulations in this section is presented in Appendix **??**. An abridgment of this section goes thus. The precept of the kinematics for a many-bodied system is introduced; we press definitions, and introduce the state partitioning scheme (); we then renew and formalize the precept of flocks' (an)isotropy (). Flock motion verification is constructed via HJI analysis () and a means for synthesizing multiple flock verification is presented ().

We locally synthesize the kinematics of agents in a manner amenable to state representation by resolving local payoff extremals, $\{\Phi_1, \cdots, \Phi_{n_f}\}$. This is a state space partition induced by an aggregation of desired collective behavior from local flocks' values $\{V_1, \cdots, V_{n_f}\}$[¶]. Suppose that the local control laws are properly coordinated, the region of the state space across which their coordinated influence might be exerted constitute a larger e.g. *manipulability volume* for a dexterous kinematic task. We now formalize definitions that will aid the modularization of the problem into manageable forms.

Definition 1. Neighbors of an Agent. *We define the neighbors $\mathcal{N}_i(t)$ of agent i at time t as the set of all agents that lie within a predefined radius, r_i.*

Definition 2. *We define a flock, F, consisting of agents labeled $\{1, 2, \cdots, n_a\}$ as a collection of agents within a phase space $(\mathcal{X}, T)$ such that all agents within the flock interact with their nearest neighbors in a topological sense.*

Remark 1. *Every agent within a flock has similar dynamics to that of its neighbor(s). Furthermore, agents travel at the same linear speed, v; the angular headings, w, however, may be different between agents, seeing we are dealing with a many-bodied system. Each agent's continuous-time dynamics, $\dot{x}^{(i)}(t)$, evolves as*

$$\begin{bmatrix} \dot{x}_1^{(i)}(t) \\ \dot{x}_2^{(i)}(t) \\ \dot{x}_3^{(i)}(t) \end{bmatrix} = \begin{bmatrix} v(t)\cos x_3^{(i)}(t) \\ v(t)\sin x_3^{(i)}(t) \\ \langle w^{(i)}(t) \rangle_r \end{bmatrix}, \tag{12}$$

$$\text{where } \langle w^{(i)}(t) \rangle_r = \frac{1}{1 + n_i(t)} \left(w^{(i)}(t) + \sum_{j \in \mathcal{N}_i(t)} w_j(t) \right) \tag{13}$$

for agents $i = \{1, 2, 3, ..., n_a\}$, where t is the continuous-time index, $n_i(t)$ is the number of agent i's neighbors at time t, $\mathcal{N}_i(t)$ denotes the sets of labels of agent i's neighbors at time t, and $\langle w^{(i)}(t) \rangle_r$ is the average orientation of agent i and its neighbors at time t. Note that for a game where all agents share the same constant linear speed and heading, (13) reduces to the dynamics of a Dubins' vehicle in absolute coordinates with $-\pi \leq w^{(i)}(t) < \pi$. The averaging over the degrees of freedom of other agents in (13) is consistent with the mean field theory, whereby the effect of all other agents on any one agent is an approximation of a single averaged influence.

Definition 3. Payoff of a Flock. *To every flock F_j (with a finite number of agents n_a) within a murmuration, $j = \{1, 2, \cdots, n_f\}$, we associate a payoff, $\boldsymbol{\Phi}_j$, that is the union of all respective agent's payoffs for expressing the outcome of a desired kinematic behavior.*

Given the recent results in robust numerical optimization of level sets of late, the last point is more of an axiom, than an assumption (see (Adalsteinsson and Sethian 1999; Saye and Sethian 2011; Zaitzeff et al. 2019; Karnakov et al. 2021)). Viscosity solutions provide a particular means of finding a unique solution with a clear interpretation in terms of the generalized optimal control problem, even in the presence of stochastic perturbations. Each agent within a flock interacts with a fixed number of neighbors, n_c, within a fixed topological range, r_c. This topological range is consistent with findings in collective swarm behaviors and it reinforces *group cohesion* (Ballerini et al. 2008). However, we are interested in *robust group cohesion* in reachability analysis. Therefore, we let a pursuer, P, with a worst-possible disturbance attack the flock, and we take it that flocks of agents constitute an evading player, E. Returning to (13), for a single flock, we now provide a sketch for the HJI formulation for a heading consensus problem.

Framework for Separated Payoffs.

Suppose that a murmuration's global heading is predetermined and each agent i within each flock, F_j, $(j =$

[‡] Note that the target set, $\mathcal{L}_0$, is a closed subset of $\mathbb{R}^n$ and is in the closure of Ω.

[§] The (backward) horizon, $-T$ is negative for $T > 0$.

[¶] Let the cursory reader understand that we use the concept of a flock loosely. The value function could represent a palette of composed value functions whose extremals resolve local behaviors we would like to emerge over separated local regions of the state space of dexterous drone acrobatics (Kaufmann et al. 2020), a robot balls juggling task (Burridge, R R and Rizzi, A A and Koditschek, D E 1999) or any parallel task domain verification problem.

$\{1, \cdots, n_f\})$ in the murmuration has a constant linear velocity, v^i. An agent's orientation is its control input, given by the average of its own orientation and that of its neighbors. Instead of metric distance interaction rules that make agents very vulnerable to predators (Ballerini et al. 2008), we resort to a topological interaction rule. With metric distance rules, we will have to formulate the breaking apart of value functions that encode a consensus heading problem in order to resolve the extrema of multiple payoffs; which is typically what we want to prevent in real-world autonomous tasks.

What constitutes an agent's neighbors are computed based on empirical findings and studies from the lateral vision of birds and fishes (Ballerini et al. 2008; Jadbabaie et al. 2003; Helbing et al. 2000) that provide insights into their anisotropic kinematic density and structure. Importantly, starlings' lateral visual axes and their lack of a rear sector reinforces their lack of nearest neighbors in the front-rear direction. As such, this enables them to maintain a tight density and robust heading during formation and flight. The delineation of an agent's nearest neighbors is given in Algorithm **??**. On lines **??** and **??** of Algorithm **??**, cohesion is reinforced by leveraging the observations above. While the neighbor updates for an agent involve an $O(n^2)$ algorithm in Algorithm **??**, we are merely dealing with $6-7$ agents at a time in a local flock – making the computational cost measly.

Each agent within a flock F_j interacts with a fixed number of neighbors, n_c, within a fixed topological range, r_c. The topological range can be set as the distance between the labels of agents in a flock. This topological range is consistent with findings in collective swarm behaviors and it reinforces *group cohesion* (Ballerini et al. 2008). However, we are interested in *robust group cohesion* in reachability analysis. Therefore, we let a pursuer, P, with a worst-possible disturbance attack the flock, and we take it that flocks of agents constitute an evading player, E.

Global Isotropy via Local Anisotropy.

Structural anisotropy is not merely an effect of a preferential velocity in animal flocking kinematics but rather an explicit effect of the anisotropic interaction character itself: agents choose a mutual position on the state space in order to maximize the sensitivity to changes in heading and speed of neighbors as the neighbors' anisotropy is optimized via vision-based collision avoidance characteristically unrelated to the eye's structure (Ballerini et al. 2008).

To reinforce robust group cohesion in local flocks, we randomly simulate a pursuer P_j against an evading agent in every flock F_j so that one agent is always relative coordinates with P^j. In this specialized case, the E and P's speeds and maximum turn radii are equal: if both players start the game with the same initial velocity and orientation, the relative equations of motion show that E can mimic P's strategy by forever keeping the starting radial separation. As such, the *barrier* is closed and *the central theme in this game of kind is to determine the surface* (Merz 1972). We defer a thorough analysis of the nature of the surface to a future work.

Owing to the high-dimensionality of the state space, we cannot resolve this barrier analytically, hence we resort to numerical approximation methods – in particular, we leverage a parallel Lax-Friedrichs integration scheme (Crandall et al. 1984) which we implement in Cupy (Nishino et al. 2017) in order to provide a *consistent* and *monotone* solution to the Hamiltonians of the flocks. The assembly in the large of these respective Hamiltonians, and hence numerically robust solutions to the variational backward reeachability problem is resolved with a Voronoi tesselation of the zero-level sets of the boundaries of the flocks.

Therefore, for an agent i within a flock with index j in a murmuration, the equations of motion under attack from a predator p (see **??**) in relative coordinates is

$$\begin{bmatrix} \dot{x}_1^{(i)_j}(t) \\ \dot{x}_2^{(i)_j}(t) \\ \dot{x}_3^{(i)_j}(t) \end{bmatrix} = \begin{bmatrix} -v_e^{(i)_j}(t) + v_p^{(j)} \cos x_3^{(i)_j}(t) + \langle w_e^{(i)_j} \rangle_r x_2^{(i)_j}(t) \\ v_p^{(i)_j}(t) \sin x_3^{(i)_j}(t) - \langle w_e^{(i)_j} \rangle_r x_1^{(i)_j}(t) \\ w_p^{(j)}(t) - \langle w_e^{(i)_j}(t) \rangle_r \end{bmatrix} \tag{14}$$

for $i = 1, \cdots, n_a$ where n_a is the number of agents within a flock, $\left(x_1^{(i)_j}(t), x_2^{(i)_j}(t)\right) \in \mathbb{R}^2$, and we have $x_3^{(i)_j}(t) \in [-\pi, +\pi)$[‖]. Read $x_1^{(i)_j}(t)$: the first component of the state of an agent i at time t which belongs to the flock j in the murmuration at time t. In absolute coordinates, the equation of motion for *free agents* is

$$\begin{bmatrix} \dot{x}_1^{(i)_j}(t) \\ \dot{x}_2^{(i)_j}(t) \\ \dot{x}_3^{(i)_j}(t) \end{bmatrix} = \begin{bmatrix} v_e^{(i)_j}(t) \cos x_3^{(i)_j}(t) \\ v_e^{(i)_j}(t) \sin x_3^{(i)_j}(t) \\ \langle w_e^{(i)_j}(t) \rangle_r \end{bmatrix}. \tag{15}$$

Flock Motion from Aggregated Value Functions.

We introduce the union operator i.e. $\cup$ below as an aggregation symbol since the respective payoffs of each agent in a flock may be implicitly or explicitly constructed[**] – when it is implicitly represented, say from a signed distance function, we shall aggregate the payoff of agents 1 and 2 as

$$\cup \{\Phi_1(x,t), \Phi_2(x,t)\} \equiv \Phi_1(x,t) \cup \Phi_2(x,t) \tag{16}$$

$$= \min(\Phi_1(x,t), \Phi_2(x,t)) \tag{17}$$

otherwise, other appropriate arithmetic or logical operation shall apply.

We assume that the *value* of a flock heading control (differential game) exists. And by an extension of Hamilton's principle of least action, the terminal motion of a flock coincide with the extremal of the payoff functional

$$V(x,t) = \inf_{\beta^{(1)} \in \mathcal{B}^{(1)}} \sup_{\boldsymbol{u}^{(1)} \in \mathcal{U}^{(1)}} g^{(1)}(x(T)) \cup \cdots \inf_{\beta^{(n_f)} \in \mathcal{B}^{(n_f)}} \sup_{\boldsymbol{u}^{(n_f)} \in \mathcal{U}^{(n_f)}} g^{(n_f)}(x(T)) \tag{18}$$

where n_f is the total number of distinct flocks in a murmuration. The resolution of this equation admits a viscosity solution to the following variational terminal HJI PDE (Mitchell et al. 2005)

$$\cup_{j=1}^{n_f} \left[\cup_{i=1}^{n_a} \left(\frac{\partial V_i}{\partial t}(x,t) + \min\left[0, H^{(i)}(x^{(i)}, V_x(x,t))\right] \right) \right] = 0. \tag{19}$$

[‖] We have multiplied the dynamics by -1 so that the extremal's resolution evolves backwards in time.

[**] In resolving the zero-level sets of HJ value functions, it is typical to represent the payoff's surface as the isocontour of some function (usually a signed distance function).

with Hamiltonian,

$$H^{(i)}(t; x^{(i)}, \boldsymbol{u}^{(i)}, \boldsymbol{v}^{(i)}, p^{(i)}) = \\ \max_{u^{(i)} \in \mathcal{U}^{(i)}} \min_{v^{(i)} \in \mathcal{V}^{(i)}} \langle f^{(i)}(t; x, \boldsymbol{u}^{(i)}, \boldsymbol{v}^{(i)}), p^{(i)} \rangle. \tag{20}$$

In swarms' collective motion, when e.g. a Peregrine Falcon attacks, immediate nearest agents change direction almost instantaneously. And because of the interdependence of the orientations of individual agents with respect to one another, all other agents respond instantaneously. Thus, we only simulate a single attack against a flock within the murmuration to realize robust cohesion.

A pursuer can attack any flock within the murmuration from a distinct surface: a P direction: this side of the surface reached after penetration in the $P - [E-]$ direction is the $P - [E]$ *side*(Isaacs 1965). We attribute the term *in the small* to determine the smooth parts of the singular surface solution when a pursuer attacks, and when they are stitched together into the total solution, we shall describe them as *in the large*. There exists at least one value $\bar{\alpha}$ of α such that if $\alpha = \bar{\alpha}$, no vector in the β-vectogram[††] penetrates the surface in the E-direction. Similar arguments can be made for $\bar{\beta}$ which prevents penetration in the P-direction. We adopt (Isaacs 1965)'s terminology and call these surfaces semi-permeable surfaces (SPS).

Throughout the game, we assume that the roles of P and E do not change, so that when capture can occur, a necessary condition to be satisfied by the saddle-point controls of the players is the Hamiltonian, $H^i(x, p)$.

Theorem 1. *For a flock, F_j, the Hamiltonian is the total energy given by a summation of the exerted energy by each agent i so that we can write the main equation or total Hamiltonian of a murmuration as*

$$H(x,p) = \max_{w_e^{(k)_j} \in [\underline{w}_e^j, \bar{w}_e^j]} \min_{w_p^{(k)_j} \in [\underline{w}_p^j, \bar{w}_p^j]} \cup_{j=1}^{n_f} \\ \left[H_a^{(k)_j}(x,p) \cup \left(\cup_{i=1}^{n_a - 1} H_f^{(i)_j}(x,p) \right) \right] \tag{21}$$

$$= \cup_{j=1}^{n_f} \left(\cup_{i=1}^{n_a-1} \left[p_1^{(i)_j} \, v^{(i)_j} \cos x_3 \right.\right. \\ \left. + p_2^{(i)_j} \, v^{(i)_j} \sin x_3 + p_3^{(i)_j} \, \langle w_e^{(i)_j} \rangle_r \right] \\ \cup \left[p_1^{(k)_j} \left(v^{(k)_j} - v^{(k)_j} \cos x_3^{(k)_j} \right) - \right. \\ p_2^{(k)_j} v^{(k)_j} \sin x_3^{(k)_j} - \underline{w}_p^j |p_3^{(k)_j}| \\ \left.\left. + \bar{w}_e^j \left| p_2^{(k)_j} x_1^{(k)_j} - p_1^{(k)_j} x_2^{(k)_j} + p_3^{(k)_j} \right| \right] \right). \tag{22}$$

where $H_a^{(k)_j}(x,p)$ is the Hamiltonian of the individual under attack by a pursuing agent, P, and $H_f^{(i)_j}(x,p)$ are the respective Hamiltonians of the free agents, $i = \{1, \cdots, n_f\}$, within an evading flock in a murmuration, and not under the direct influence of capture or attack by P; we denote by $w_e^{(i)_j}$ the heading of an evader i within a flock j and $w_p^{(j)}$ the heading of a pursuer aimed at flock j; $\underline{w}_e^{(k)_j}$ is the orientation that corresponds to the orientation of the agent with minimum turn radius among all the neighbors of agent k, inclusive of agent k at time t; similarly, $\bar{w}_e^{(k)_j}$ is the maximum orientation among all of the orientation of agent k's neighbors.

Corollary 1. *For the special case where the linear speeds of the evading agents and pursuer are equal i.e. $v_e^{(i)_j}(t) = v_p(t) = +1m/s$, we have the Hamiltonian as*

$$H(x,p) = \cup_{j=1}^{n_f} \left(\cup_{i=1}^{n_a-1} \left[p_1^{(i)_j} \, \cos x_3 + p_2^{(i)_j} \, \sin x_3 \right.\right. \\ \left. + p_3^{(i)_j} \, \langle w_e^{(i)_j} \rangle_r \right] \cup \left[p_1^{(k)_j} \left(1 - \cos x_3^{(k)_j} \right) - p_2^{(k)_j} \sin x_3^{(k)_j} \right. \\ \left.\left. - \underline{w}_p^j |p_3^{(k)_j}| + \bar{w}_e^j \left| p_2^{(k)_j} x_1^{(k)_j} - p_1^{(k)_j} x_2^{(k)_j} + p_3^{(k)_j} \right| \right] \right). \tag{23}$$

We adopt the essentially non-oscillatory Lax-Friedrichs scheme of (Osher and Shu 1991; Crandall and Lions 1984a) in resolving (23). Denote by (x, y, z) a generic point in $\mathbb{R}^3$ so that given mesh sizes Δx, Δy, Δz, $\Delta t > 0$, letters u, v, w will represent functions on the x, y, z lattice $\Delta = \{(x_i, y_j, z_k) : i, j, k \in \mathbb{Z}\}$. We define the numerical monotone flux, $\hat{H}^{(i)_j}(\cdot)$, of $H_j^{(i)}(\cdot)$ as

$$\hat{H}^{(i)_j}(u^+, u^-, v^+, v^-, w^+, w^-) = \\ H^{(i)_j} \left(\frac{u^+ + u^-}{2}, \frac{v^+ + v^-}{2}, \frac{w^+ + w^-}{2} \right) \\ - \frac{1}{2} \left[\alpha_x^{(i)_j} \left(u^+ - u^- \right) + \alpha_y^{(i)_j} \left(v^+ - v^- \right) + \alpha_z^{(i)_j} \left(w^+ - w^- \right) \right] \tag{24}$$

where

$$\alpha_x^{(i)_j} = \max_{\substack{a \le u \le b \\ c \le v \le d \\ e \le w \le f}} |H_u^{(i)_j}(\cdot)|, \; \alpha_y^{(i)_j} = \max_{\substack{a \le u \le b \\ c \le v \le d \\ e \le w \le f}} |H_v^{(i)_j}(\cdot)|, \\ \alpha_z^{(i)_j} = \max_{\substack{a \le u \le b \\ c \le v \le d \\ e \le w \le f}} |H_w^{(i)_j}(\cdot)| \tag{25}$$

are the dissipation coefficients, controlling the level of numerical viscosity in order to realize a stable solution that is physically realistic (Crandall and Lions 1984a). Here, the subscripts of H are the partial derivatives w.r.t the subscript variable, and the flux, $\hat{H}(\cdot)$ is monotone for $a \le u^{\pm} \le b, c \le v^{\pm} \le d, e \le w^{\pm} \le f$. We adopt the total variation diminishing Runge-Kutta scheme of (Osher and Shu 1987) in efficiently calculating essentially non-oscillating upwinding finite difference gradients of $H(\cdot)$.

ComplexBRAT by Voronoi Tesselaton of Local ϵ-BRAT Interfaces.

The method we propose here is inspired by the algorithmic notions of robust, self-organizing emergent "behaviors" where efficiency and consistency is important when considering the interconnection between moving interfaces (Saye and Sethian 2011). In our case, the physics of the local interface of the flocks that constitute a murmuration possesses topological complexity arising from local value function boundaries that evolve temporally e.g. via intersection or destruction of interfaces as a result of physical phenomena changes such as vacuole, splitting, or flash expansion inherent in starlings murmurations (cf. **??**).

[††] A $\beta-$vectogram is the resulting state space when a the strategy β is applied in computing the optimal control law for an agent.

Suppose that the boundaries between two flocks F_j, F_k is a closed hypersurface, moving through time i.e. $\Gamma_{jk}(t=0) \in \mathbb{R}^N$ with speed v_j as given in (29). We can solve the internally generated level set equation (29) to obtain the capture surface of respective flocks (see subsection). Given the value function aggregation scheme (cf. 19), we are interested in solving for the externally generated velocity field, v_{ext}, induced by a flock's kinematics in light of (14) so that v_j in (14) is now a parameter for every level set, instead of just the level set of the interface alone. We call v_{ext} the external velocity, so that at the zero level set, we have

$$v_{ext}^j = v_j \text{ when } V_j = 0. \tag{26}$$

The key idea here is that we stitch the interfaces $\Gamma_{jk}, \Gamma_{kl}, \Gamma_{lm}$, by leveraging motion involving mean curvature (Osher and Fedkiw 2004, §4.1) in systems characterized by multi-phase kinetics. Therefore, a gradient descent on the energies of the respective flocks at the zero level set can be computed from

$$\cup_{i=1}^{n_f} V_i = \sum_{j \neq k} \gamma_{jk} \Upsilon(\Gamma_{jk}), \text{ for } (j,k) \in \{i \mid i = 1, \cdots, n_f\}$$
$$\text{where } V_i \in \Omega, \tag{27}$$

Υ is the area covered by the multi-flock interface, $V_i \cap V_j = (\partial V_i) \cap (\partial V_j)$ for $i \neq j$, $V_i \cap V_i = (\partial V_i) \cap (\partial V_i) \equiv \emptyset$. In addition, we require the surface tension, γ_{jk}, to be positive so that the interfaces shrink as the level set equation evolves over time; we do this by imposing the following triangle inequality $\gamma_{jk} + \gamma_{kl} \geq \gamma_{jl}$ for distinct j,k,l in order to assure that (27) is well-posed (Zaitzeff et al. 2019). Where interacting flocks share a boundary, we characterize such higher order junctions by triple hypersurfaces, by tracking the ϵ-BRATs for an $\epsilon > 0$. The kinematics of these ϵ-BRATs under attack by predators constitute the evolution of the event-driven behavior of murmurations that swoop, swirl, or whirl in order to evade capture.

The speed of the interface, at a point $x \in \Gamma_{jk}$ a distance from a triple junction, is given by

$$v_N(x) = \gamma_{jk} \mathcal{K}_{jk}(x) \tag{28}$$

for a curvature $\mathcal{K}_{jk}(x) = \mathcal{K}_{kj}(x)$, and a normal speed $v_N(x)$. We implicitly initialize the payoff, Φ_j^i of each agent labeled $p \in \{1, \cdots n_a\}$ within every flock, F_j, $\forall\, j = 1, \cdots, n_f$ as a signed distance function $d_{\Phi_j^i}(x)$ to a phase (set) in $\Phi_j^i \in \Omega$ so that it yields an Euclidean distance to the boundary $\partial \Phi_j^i$ whose sign bit is an indicator function, signifying that if a point $x \in \mathbb{R}^n$ is inside or outside Φ_j^i i.e.

$$d_{\Phi_j^i}(x) = \begin{cases} \inf_{\boldsymbol{z}^{(i)} \in \partial \Phi_j^i} \mid x - \boldsymbol{z}^{(i)} \mid, & x \in \Phi_j^i, \\ -\inf_{\boldsymbol{z}^{(i)} \in \partial \Phi_j^i} \mid x - \boldsymbol{z}^{(i)} \mid, & x \notin \Phi_j^i, \end{cases} \tag{29}$$

so that each agent's initial position is uniquely represented on the overall vectogram based on the value of $\boldsymbol{z}^{(i)}$. In order to maintain a consistent level set representation for each payoff, (e.g. when flocks split, expand, or spread out), the structure of interface must be maintained as time evolves. We follow (Adalsteinsson and Sethian 1999)'s construction and write the level set equation as

$$v_{ext}^j \cdot \nabla V_j = 0, \text{ where } v_{ext}^j = \frac{\nabla \Phi}{|\nabla \Phi|}, \tag{30}$$

so that the level set function V_j remains the signed distance function as time evolves. When the level set functions must be evolved concurrently, we reparameterize the level set equation with an unsigned distance function as a union of an $\epsilon > 0$ super-level sets of the respective flocks

$$\frac{\partial \Phi}{\partial t}(x,t) - \nabla \cdot v_{ext}^j \, |\nabla \Phi(x,t)| = \epsilon. \tag{31}$$

That is the level set corresponding to the interface is now a neighbor of nearby level sets; this makes the motion of the zero level set that corresponds to the interface (29) is surrounded by the motion of nearby level sets. Similar to (Saye and Sethian 2011), we define Voronoi interface Γ_V as the set of of all points x that are equidistant to at least two different ϵ-level sets belonging to different flocks, and no closer to any other ϵ-level set i.e.

$$\Gamma_V = \{x \in \Omega : \exists i \neq j\}$$
$$\text{such that } d(x, \Gamma_{\epsilon,i}) = d(x, \Gamma_{\epsilon,j}) \leq d(x, \Gamma_{\epsilon,k}) \,\forall\, k \neq i,j, \tag{32}$$

where $\Gamma_{\epsilon,i}$ is the ϵ-level set corresponding to a flock i.

Safe Trajectory Optimization: Generalized Nash Equilibria

Experiments.

At issue is a family of games with different target sets for local flocks that on the whole constitute a murmuration. Every agent's target position is initialized as

$$\rho_j \begin{bmatrix} r_c \cos(\frac{i\pi}{4}), & r_c \sin(\frac{i\pi}{4}), & h + i\,\delta h \end{bmatrix}^T$$
$$\forall\, i \in \{1, \cdots n_a\}, \forall\, j \in \{1, \cdots n_f\}. \tag{33}$$

Here, ρ_j is a scaling factor that ensures adequate *inter-flock separation* on a grid, h, δh are appropriately problem-dependent parameters, and r_c is a collision avoidance radius. The set of grid points for which the states of (33) is defined are those point set for which $d_{\Phi_j^i}(x)$ is fixed and $\Omega = \{\text{all grid points}\}$. We set $d_{\Phi_j^i}(x) = \text{sgn}(V_j^i(x))$ for all $x \in \Omega$. Fig. 1 denote the representation of the payoffs of certain agents that constitute a flock. They are constructed from the signed distance function from all points on the grid to an interface in the spirit of the foregoing.

An adaptive allocation rule for robust cohesion lets P randomly aim against an agent within an evading flock in every iteration of the game (see **??**) – since when hunting for a prey, an originally targeted prey may evade P. The domain in which we calculate the BRT of the agent under attack in relative coordinates w.r.t a pursuer, and that of the other agents (within a flock) in absolute coordinates are respectively

$$\bar{\Omega}_{rel} = \mathbb{R}^2 \times \mathcal{S}^1, \quad \bar{\Omega}_{abs} = \{\cup_{p=1}^{n_a - 1} \bar{\Omega}_{abs}^p | \bar{\Omega}_{abs}^p \in \mathbb{R}^2 \times \mathcal{S}^1\}. \tag{34}$$

It follows that the domain that constitutes the BRAT for a flock F_j is

$$\bar{\Omega}_{flk}^j = \{\bar{\Omega}_{rel}^j \cup \bar{\Omega}_{abs}^j\}, \tag{35}$$

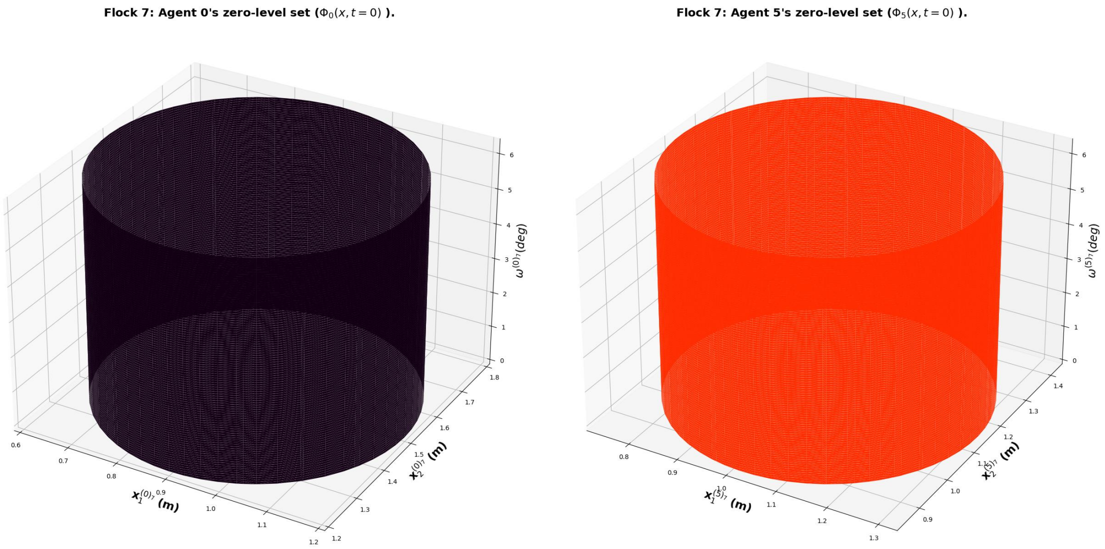


Figure 1. Implicit representation of the payoffs for agents 2 and 5 within flock 7.

and the domain that constitutes the BRAT of a murmuration $F_j, \cdots, F_{n_f}$ is

$$\bar{\Omega}_{murmur} = \{\bar{\Omega}^{j}_{flk} \cup \bar{\Omega}^{j+1}_{flk} \cup \cdots \cup \bar{\Omega}^{n_f}_{flk}\}. \tag{36}$$

Since the orientations of neighboring agents are averaged throughout a flock, the information is inevitably propagated across the entire flock. Note that the above equations imply that the cardinality of all agents within a flock is $[n_a]$ and the cardinality of all agents within a murmuration is $[n_f]$. We define the payoff for a flock F_j as the union of the payoff of every individual agent that constitute it (that is, it is the union of the respective payoffs as shown in Fig. 1) i.e.

$$\Phi_j = \Phi_1 \cup \Phi_2 \cup \cdots \cup \Phi_{n_a-1} \cup \Phi_{n_a}, \tag{37}$$

where a Φ solves the level set PDE (Sethian 2000) in the form of the *unsigned distance function* of (31). The backward reach-avoid tubes we aim to compute constitute the states' set from which the pursuer can drive the evader into the target set

$$\mathcal{L}_0 = \left\{x \in \bar{\Omega} | \left(\sqrt{x_1^{(1)2} + x_2^{(1)2}} - r_c^{(1)}\right) \cup \left(\sqrt{x_1^{(2)2} + x_2^{(2)2}} - r_c^{(2)}\right) \cup \cdots \right.$$
$$\left. \cup \left(\sqrt{x_1^{(n_a)2} + x_2^{(n_a)2}} - r_c^{(n_a)}\right)\right\} \tag{38}$$

where superscripts in parentheses denote the label of an agent, so that in a 3-D, $\mathcal{L}_0$ is akin to an uneven cylinder (see left inset of Fig. 2). The interface is the union of the zero-level set of the payoffs of the individual agents that constitute the flock (see Fig. 1). Each flock's zero-level set is distinct because the target set of its agents belong to unique positions in the state space.

Results from the numerical integration of the level set equation for different flocks is depicted in Fig. 2, and more results for other flocks are included in the Appendix . Let us enquire. **Observe**: (a) Each flock's RCBRAT' surface is nonconvex; (b) The Lax-Friedrich's numerical integration scheme of the respective Hamilton-Jacobi value functions has discontinuities in the solution despite the value function being initially smooth; (c) Owing to the possibility of non-unique solutions to each initial value problem, the weighted essentially nonoscillatory entropy scheme we adopted helps in picking out "physically" relevant solutions to (22).

Conclusion.

We have proposed an Hamilton-Jacobi-Isaacs systems verification scheme, based on Hamilton-Jacobi's reachability theory for constructing backward reachable tubes for a complex system with structural local behaviors that is characterized by topological nearest neighbor rules. These local spatio-temporal dynamics whose local interactions constitute a collective behavior. Using the key idea that the total energy within every subsystem is an aggregate of the respective energies of its individual agents, we have formulated a theorem for constructing the local Hamiltonians as well as value functionals.

Under the assumptions that we have (i) constant linear velocity among agents; (ii) each agent's orientation serves as the control input; (iii) intra-flock agent interaction occurs within unique and distinct state space manifolds; and by agents maneuvering their direction, a kinematic alignment is obtained with other flocks; and (iv) inter-flock interaction occurs when a pursuer is within a threshold of capturing any agent within the murmuration, we have presented a numerical input/state-constraint preserving scheme utilizing a time-constrained HJI formulation in a backward reachability setting.

References

Adalsteinsson D and Sethian JA (1999) The Fast Construction Of Extension Velocities In Level Set Methods. *Journal of Computational Physics* 148: 2–22.

Bajcsy A, Bansal S, Bronstein E, Tolani V and Tomlin CJ (2019) An Efficient Reachability-based Framework for Provably Safe Autonomous Navigation in Unknown Environments. In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, pp. 1758–1765.

Balcázar N, Lehmkuhl O, Rigola J and Oliva A (2015) A Multiple Marker Level-set Method for Simulation of Deformable Fluid

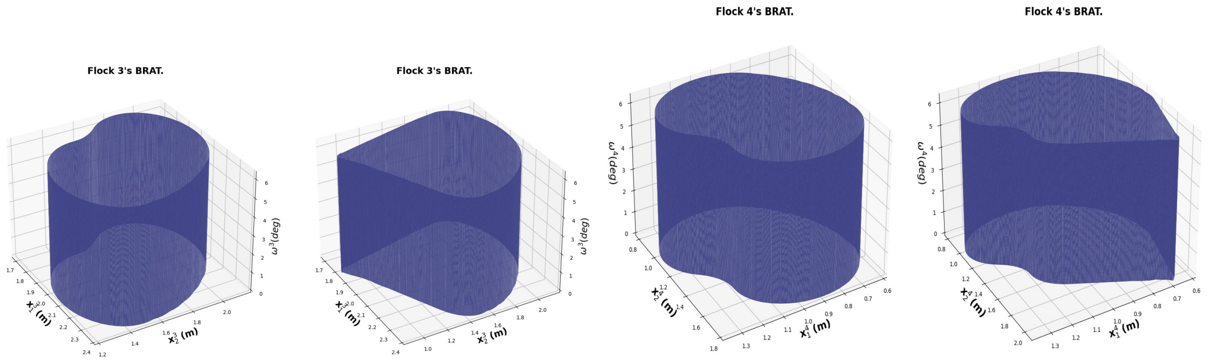


Figure 2. Left: Interface or zero-level set for two example flocks. Right: Interfaces (zero-level set) of the evading flocks under attack from a pursuer at the end of the respective Lax-Friedrichs' integration scheme. Metric reach radius=$0.2m$, avoid Radius=$0.2m$. More results are included in Appendix .

Particles. *International Journal of Multiphase Flow* 74: 125–142.

Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, Giardina I, Lecomte V, Orlandi A, Parisi G, Procaccini A, Viale M and Zdravkovic V (2008) interaction Ruling Animal Collective Behavior Depends On Topological Rather Than Metric Distance: Evidence From A Field Study. *Proceedings of the National Academy of Sciences* 105(4): 1232–1237. DOI:10.1073/pnas.0711437105. URL https://www.pnas.org/content/105/4/1232.

Bansal S and Tomlin CJ (????) DeepReach: A Deep Learning Approach to High-Dimensional Reachability .

Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M and Walczak AM (2012) Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Sciences* 109(13): 4786–4791.

Burridge, R R and Rizzi, A A and Koditschek, D E (1999) Sequential Composition of Dynamically Dexterous Robot Behaviors. Technical Report 6.

Cavagna A, Cimarelli A, Giardina I, Parisi G, Santagati R, Stefanini F and Viale M (2010) Scale-ffee correlations in starling flocks. *Proceedings of the National Academy of Sciences* 107(26): 11865–11870.

Clarke FH, Ledyaev YS, Stern RJ and Wolenski PR (2008) *Nonsmooth Analysis and Control Theory*, volume 178. Springer Science & Business Media.

Crandall M and Majda A (1980) The method of fractional steps for conservation laws. *Numerische Mathematik* 34(3): 285–314.

Crandall MG, Evans LC and Lions PL (1984) Some Properties of Viscosity Solutions of Hamilton-Jacobi Equations. *Transactions of the American Mathematical Society* 282(2): 487.

Crandall MG and Lions PL (1983) Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society* 277(1): 1–42.

Crandall MG and Lions PL (1984a) Two Approximations of Solutions of Hamilton-Jacobi Equations. *Mathematics of Computation* 43(167): 1.

Crandall MG and Lions PL (1984b) Two approximations of solutions of hamilton-jacobi equations. *Mathematics of computation* 43(167): 1–19.

Dubins LE (1957) On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents. *American Journal of mathematics* 79(3): 497–516.

Evans L and Souganidis PE (1984) Differential Games And Representation Formulas For Solutions Of Hamilton-Jacobi-Isaacs Equations. *Indiana Univ. Math. J* 33(5): 773–797.

Fisac JF, Akametalu AK, Zeilinger MN, Kaynama S, Gillula J and Tomlin CJ (2019a) A General Safety Framework for Learning-Based Control in Uncertain Robotic Systems. *IEEE Transactions on Automatic Control* 64(7): 2737–2752. DOI: 10.1109/TAC.2018.2876389.

Fisac JF, Lugovoy NF, Rubies-Royo V, Ghosh S and Tomlin CJ (2019b) Bridging hamilton-jacobi safety analysis and reinforcement learning. *Proceedings - IEEE International Conference on Robotics and Automation* 2019-May: 8550–8556. DOI:10.1109/ICRA.2019.8794107.

Gerdts M, Karrenberg S, Müller-Beßler B and Stock G (2009) Generating Locally Optimal Trajectories for an Automatically Driven Car. *Optimization and Engineering* 10(4): 439–463.

Haiken M (2021) These birds flock in mesmerizing swarms of thousands—but why is still a mystery. URL https://tinyurl.com/4973byey.

Helbing D, Farkas I and Vicsek T (2000) Simulating dynamical features of escape panic. *Nature* 407(6803): 487–490.

Herbert S, Choi JJ, Sanjeev S, Gibson M, Sreenath K and Tomlin CJ (2021) Scalable learning of safety guarantees for autonomous systems using hamilton-jacobi reachability. *arXiv preprint arXiv:2101.05916* .

Herbert SL, Chen M, Han S, Bansal S, Fisac JF and Tomlin CJ (2017) Fastrack: A modular framework for fast and guaranteed safe motion planning. In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, pp. 1517–1522.

Hobbs KL, Mote ML, Abate MC, Coogan SD and Feron EM (2023) Runtime assurance for safety-critical systems. *IEEE Control Systems Magazine* 43: 28–65.

Holland J (1975) Adaptation in Natural and Artificial Systems. *Ann Arbor* .

Isaacs R (1965) Differential games: A mathematical theory with applications to warfare and pursuit, control and optimization. *Kreiger, Huntigton, NY* .

Jadbabaie A, Lin J and Morse AS (2003) Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on automatic control* 48(6): 988–1001.

Kappen HJ (2005) Linear theory for control of nonlinear stochastic systems. *Physical review letters* 95(20): 200201.

Karnakov P, Litvinov S and Koumoutsakos P (2021) Computing Foaming Flows Across Scales: From Breaking Waves to

Microfluidics. *arXiv preprint arXiv:2103.01513* .

Kaufmann E, Loquercio A, Ranftl R, Müller M, Koltun V and Scaramuzza D (2020) Deep Drone Acrobatics. *arXiv preprint arXiv:2006.05768* .

Kim Y, Lai MC and Peskin CS (2010) Numerical Simulations of Two-Dimensional Foam by the Immersed Boundary Method. *Journal of Computational Physics* 229(13): 5194–5207.

Kirkpatrick S, Gelatt Jr CD and Vecchi MP (1983) Optimization by Simulated Annealing. *Science* 220(4598): 671–680.

Kobilarov M (2012) Cross-entropy randomized motion planning. In: *Robotics: Science and Systems*, volume 7. pp. 153–160.

LaValle SM (2006) *Planning algorithms*. Cambridge university press.

Leonid Koralov YGS (2007) *Theory of Probability and Random Processes*. Springer.

Li T, Peng G, Zhu Q and Başar T (2022) The Confluence of Networks, Games, and Learning a Game-Theoretic Framework for Multiagent Decision Making Over Networks. *IEEE Control Systems Magazine* 42(4).

Lions PL (1982) *Generalized solutions of Hamilton-Jacobi equations*, volume 69. London Pitman.

Lygeros J (2004) On reachability and minimum cost optimal control. *Automatica* 40(6): 917–927.

Merz A (1972) The game of two identical cars. *Journal of Optimization Theory and Applications* 9(5): 324–343.

Mitchell I (2001) Games of two identical vehicles. *Dept. Aeronautics and Astronautics, Stanford Univ.* (July): 1–29.

Mitchell I (2020) A Robust Controlled Backward Reach Tube with (Almost) Analytic Solution for Two Dubins Cars. *EPiC Series in Computing* 74: 242–258.

Mitchell IM, Bayen AM and Tomlin CJ (2005) A Time-Dependent Hamilton-Jacobi Formulation of Reachable Sets for Continuous Dynamic Games. *IEEE Transactions on Automatic Control* 50(7): 947–957. DOI:10.1109/TAC.2005.851439.

Nishino R, Loomis C and Hido S (2017) Cupy: A numpy-compatible library for nvidia gpu calculations. *31st confernce on neural information processing systems* 151.

Osher S and Fedkiw R (2004) Level Set Methods and Dynamic Implicit Surfaces. *Applied Mechanics Reviews* 57(3): B15–B15.

Osher S and Sethian JA (1988) Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of computational physics* 79(1): 12–49.

Osher S and Shu CW (1987) Efficient Implementation of Essentially Non-Oscillatory Shock Capturing Schemes. Technical report, NASA Langley Research Center, Hampton, Virginia.

Osher S and Shu CW (1991) High-Order Essentially Nonoscillatory Schemes for Hamilton-Jacobi Equations. *SIAM Journal of Numerical Analysis* 28(4): 907–922.

Pavliotis GA (2014) *Stochastic Processes and Applications*. Springer.

Saravanos AD, Aoyama Y, Zhu H and Theodorou EA (2022) Distributed Differential Dynamic Programming Architectures for Large-Scale Multi-Agent Control. *arXiv preprint arXiv:2207.13255* .

Saye RI and Sethian JA (2011) The Voronoi Implicit Interface Method for Computing Multiphase Physics. *Proceedings of the National Academy of Sciences of the United States of America* 108(49): 19498–19503.

Sethian JA (1987) Numerical Methods for Propagating Fronts. In: *Variational methods for free surface interfaces*. Springer, pp. 155–164.

Sethian JA (1996) A Fast Marching Level Set Method For Monotonically Advancing Fronts. *Proceedings of the National Academy of Sciences* 93(4): 1591–1595.

Sethian JA (2000) Level Set Methods And Fast Marching Methods: Evolving Interfaces In Computational Geometry, Fluid Mechanics, Computer Vision, And Materials Science. *Robotica* 18(1): 89–92.

Song Y, Steinweg M, Kaufmann E and Scaramuzza D (2021) Autonomous Drone Racing with Deep Reinforcement Learning. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 1205–1212.

Theodorou E, Buchli J and Schaal S (2010) A Generalized Path Integral Control Approach to Reinforcement Learning. *The Journal of Machine Learning Research* 11: 3137–3181.

Torrente G, Kaufmann E, Föhn P and Scaramuzza D (2021) Data-driven MPC for Quadrotors. *IEEE Robotics and Automation Letters* 6(2): 3769–3776.

Tsitsiklis JN (1995) Globally Optimal Trajectories. *IEEE Transactions on Automatic Control* 40(9): 1528–1538.

Vicsek T, Czirók A, Ben-Jacob E, Cohen I and Shochet O (1995) Novel type of phase transition in a system of self-driven particles. *Physical review letters* 75(6): 1226.

Vinitsky E, Lichtlé N, Yang X, Amos B and Foerster J (2022) Nocturne: a scalable driving benchmark for bringing multi-agent learning one step closer to the real world. *arXiv preprint arXiv:2206.09889* URL https://arxiv.org/abs/2206.09889.

Williams G, Aldrich A and Theodorou EA (2017) Model predictive path integral control: From theory to parallel computation. *Journal of Guidance, Control, and Dynamics* 40(2): 344–357.

Williams G, Drews P, Goldfain B, Rehg JM and Theodorou EA (2016) Aggressive Driving With Model Predictive Path Integral Control. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 1433–1440.

Zaitzeff A, Esedoglu S and Garikipati K (2019) On the Voronoi Implicit Interface Method. *SIAM Journal on Scientific Computing* 41(4): A2407–A2429.

Hamiltonian of a Murmuration.

In this appendix, we provide a derivation for the Hamiltonian of a flock, and by extension, that of a murmuration. In our implementations, the zero-level set is constructed implicitly from the isocontour of a signed distance function as described in (Osher and Fedkiw 2004, Chapter II). We introduce the union operator, ∪, below in lieu of the arithmetic summation symbol since the respective payoffs of each agent in a flock are implicitly initialized as signed distance functions on the state space. It is trivial to extend these results to other arithmetic or Boolean operators depending on different task domains.

Recall from (22) that the total Hamiltonian of a flock is a union of the mechanical energy of the free agents in a flock and the individual under attack i.e.

$$H(x,p) = \max_{w_e^{(k)_j}\in[\underline{w}_e^j,\bar{w}_e^j]} \min_{w_p^{(k)_j}\in[\underline{w}_p^j,\bar{w}_p^j]} \cup_{j=1}^{n_f} \left[H_a^{(k)_j}(x,p) \cup \left(\cup_{i=1}^{n_a-1} H_f^{(i)_j}(x,p)\right)\right] \tag{39}$$

Proof of Theorem 1. We write the Hamiltonian of the free agents in absolute coordinates and the Hamiltonian of the agent under attack in relative coordinates with respect to the pursuer. A flock's Hamiltonian is Hamiltonian of the free agents is the aggregation of all the mechanical energy in the system in absolute coordinates i.e.

$$\cup_{i=1}^{n_a-1} H_f^{(i)_j}(x,p) = \cup_{i=1}^{n_a-1}\begin{bmatrix}p_1^{(i)_j} & p_3^{(i)_j} & p_3^{(i)_j}\end{bmatrix} \begin{bmatrix} v^{(i)_j}\cos x_3 \\ v^{(i)_j}\sin x_3 \\ \langle w_e^{(i)_j}\rangle_r \end{bmatrix} \tag{40}$$

where we have again dropped the time arguments for convenience. It follows that

$$\cup_{i=1}^{n_a-1} H_f^{(i)_j}(x,p) = \\ \cup_{i=1}^{n_a-1}\left[p_1^{(i)_j}\, v^{(i)_j}\cos x_3 + p_3^{(i)_j}\, v^{(i)_j}\sin x_3 + p_3^{(i)_j}\,\langle w_e^{(i)_j}\rangle_r\right]. \tag{41}$$

Equation (22) can be re-written as

$$H_a^{(k)_j}(x,p) = \\ -\left(\max_{w_e^{(k)_j}\in[\underline{w}_e^j,\bar{w}_e^j]} \min_{w_p^{(k)_j}\in[\underline{w}_p^j,\bar{w}_p^j]} \begin{bmatrix}p_1^{(k)_j}(t) & p_2^{(k)_j}(t) & p_3^{(k)_j}(t)\end{bmatrix} \begin{bmatrix} -v_e^{(k)_j}(t) + v_p^{(j)}\cos x_3^{(k)_j}(t) + \langle w_e^{(k)_j}\rangle_r(t) x_2^{(k)_j}(t) \\ v_p^j(t)\sin x_3^{(k)_j}(t) - \langle w_e^{(k)_j}\rangle_r(t) x_1^{(k)_j}(t) \\ w_p^j(t) - \langle w_e^{(k)_j}(t)\rangle_r \end{bmatrix}\right), \tag{42}$$

where $p_l^{(k)_j}(t)\mid_{l=1,2,3}$ are the adjoint vectors (Merz 1972). For the pursuer, its minimum and maximum turn rates are fixed so that we have $\underline{w}_p^j$ as the minimum turn bound of the pursuing vehicle, and $\bar{w}_p^j$ is the maximum turn bound of the pursuing vehicle. Henceforth, we drop the templated time arguments for ease of readability. Rewriting (42), we find that

$$\begin{aligned} H_a^{(k)_j}(x,p) = &-\left(\max_{w_e^{(k)_j}\in[\underline{w}_e^j,\bar{w}_e^j]} \min_{w_p^{(k)_j}\in[\underline{w}_p^j,\bar{w}_p^j]} \left[-p_1^{(k)_j} v_e^{(k)_j} + p_1^{(k)_j} v_p^j \cos x_3^{(k)_j}\right.\right. \\ &\left.\left.+ p_1^{(k)_j}\langle w_e^{(k)_j}\rangle_r x_2^{(k)_j} + p_2^{(k)_j} v_p^j \sin x_3^{(k)_j} - p_2^{(k)_j}\langle w_e^{(k)_j}\rangle_r x_1^{(i)_j} + p_3^{(k)_j}\left(w_p^j - \langle w_e^{(k)}\rangle_r\right)\right]\right), \\ = &\, p_1^{(k)_j}\left(v_e^{(k)_j} - v_p^j\cos x_3^{(k)_j}\right) - p_2^{(k)_j} v_p^j \sin x_3^{(k)_j} \\ &+ \left(\max_{\langle w_e^{(k)_j}\rangle_r\in[\underline{w}_e^j,\bar{w}_e^j]} \min_{w_p^j\in[\underline{w}_p^j,\bar{w}_p^j]} \left[\langle w_e^{(k)_j}\rangle_r\left(p_2^{(k)_j} x_1^{(k)_j} - p_1^{(k)_j} x_2^{(k)_j} + p_3^{(k)_j}\right) - p_3^{(k)_j} w_p^j\right]\right). \end{aligned} \tag{43}$$

It follows that we have from (43) that

$$H_a^{(k)_j}(x,p) = p_1^{(k)_j}\left(v_e^{(k)_j} - v_p^j\cos x_3^{(k)_j}\right) - p_2^{(k)_j} v_p^j \sin x_3^{(k)_j} - \underline{w}_p^j|p_3^{(k)_j}| \\ + \bar{w}_e^j\left|p_2^{(k)_j} x_1^{(k)_j} - p_1^{(k)_j} x_2^{(k)_j} + p_3^{(k)_j}\right| \tag{44}$$

and that

$$H_f^{(i)_j}(x,p) = \left[p_1^{(i)_j}\, v^{(i)_j}\cos x_3 + p_3^{(i)_j}\, v^{(i)_j}\sin x_3 + p_3^{(i)_j}\,\langle w_e^{(i)_j}\rangle_r\right]. \tag{45}$$

A fortiori the main equation (22) becomes

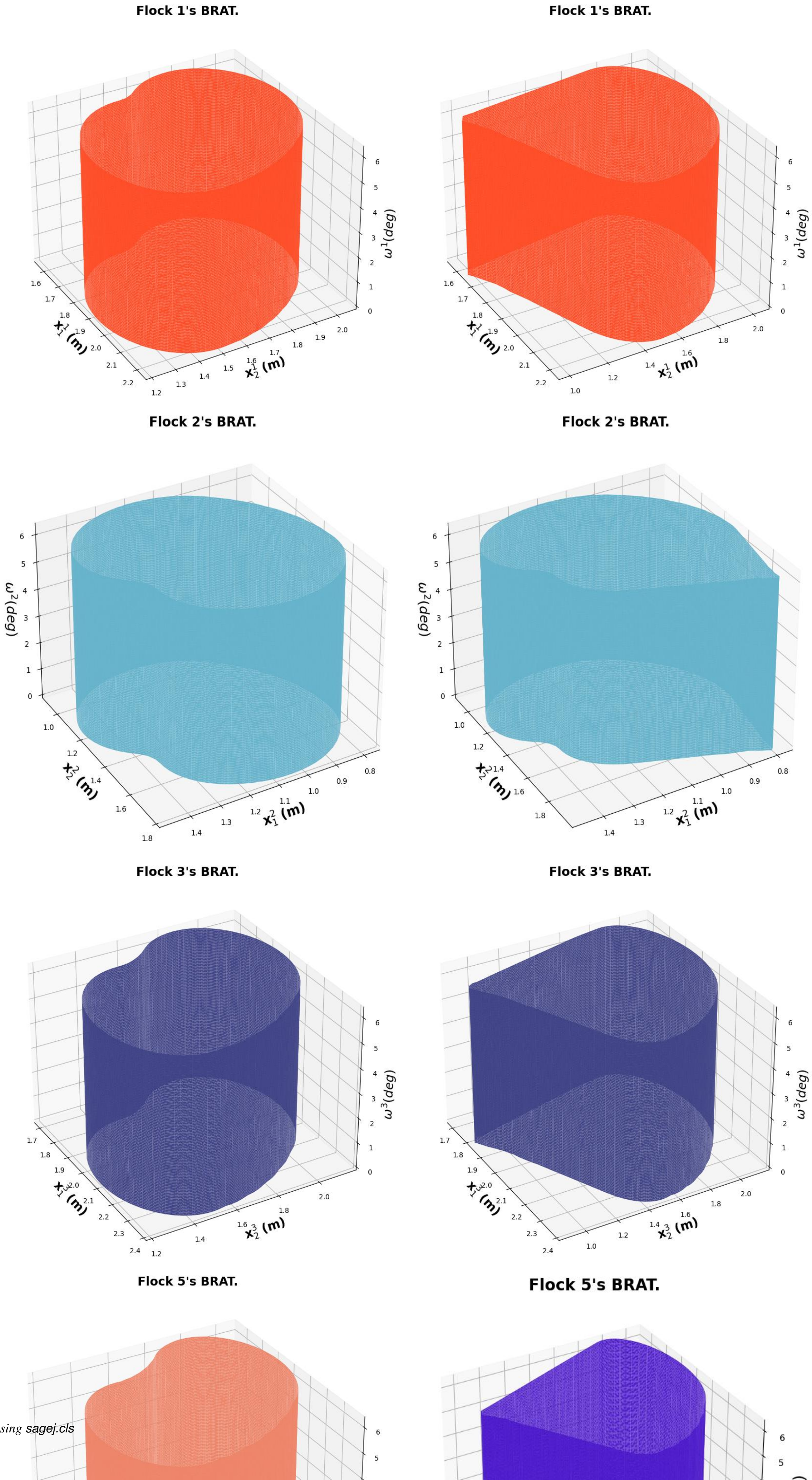
$$H(x,p) = \cup_{j=1}^{n_f}\left(\cup_{i=1}^{n_a-1}\left[p_1^{(i)_j}\, v^{(i)_j}\cos x_3 + p_2^{(i)_j}\, v^{(i)_j}\sin x_3 + p_3^{(i)_j}\,\langle w\right.\right. \\ \cup\left[p_1^{(k)_j}\left(v^{(k)_j} - v^{(k)_j}\cos x_3^{(k)_j}\right) - p_2^{(k)_j} v^{(k)_j}\sin x_3^{(k)_j} - \right. \\ \left.\left. + \bar{w}_e^j\left|p_2^{(k)_j} x_1^{(k)_j} - p_1^{(k)_j} x_2^{(k)_j} + p_3^{(k)_j}\right|\right]\right). \tag{46}$$

For the special case where the linear speeds of the evading agents and pursuer are equal i.e. $v_e^{(i)_j}(t) = v_p(t) = +1m/s$, we have a murmuration's Hamiltonian as

$$H(x,p) = \cup_{j=1}^{n_f}\left(\cup_{i=1}^{n_a-1}\left[p_1^{(i)_j}\cos x_3 + p_2^{(i)_j}\sin x_3 + p_3^{(i)_j}\,\langle w_e^{(i)_j}\rangle_r\right]\right. \\ \cup\left[p_1^{(k)_j}\left(1 - \cos x_3^{(k)_j}\right) - p_2^{(k)_j}\sin x_3^{(k)_j} - \underline{w}_p^j|p_3^{(k)_j}|\right. \\ \left.\left. + \bar{w}_e^j\left|p_2^{(k)_j} x_1^{(k)_j} - p_1^{(k)_j} x_2^{(k)_j} + p_3^{(k)_j}\right|\right]\right). \tag{47}$$

Flocks' Robustly Controllable BRATs

Note that the symmetry between non-consecutive flock labels e.g. flock 1 and flock 3's RCBRAT is because the we multiplied the initial position of a flock's state by -1.



Flock 1's BRAT.

Flock 1's BRAT.

Flock 2's BRAT.

Flock 2's BRAT.

Flock 3's BRAT.

Flock 3's BRAT.

Flock 5's BRAT.

Flock 5's BRAT.