2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Collaborative Human-Robot Object Transportation Using a Deformable Sheet

Weijian Zhang, Charlie Street, Masoumeh Mansouri

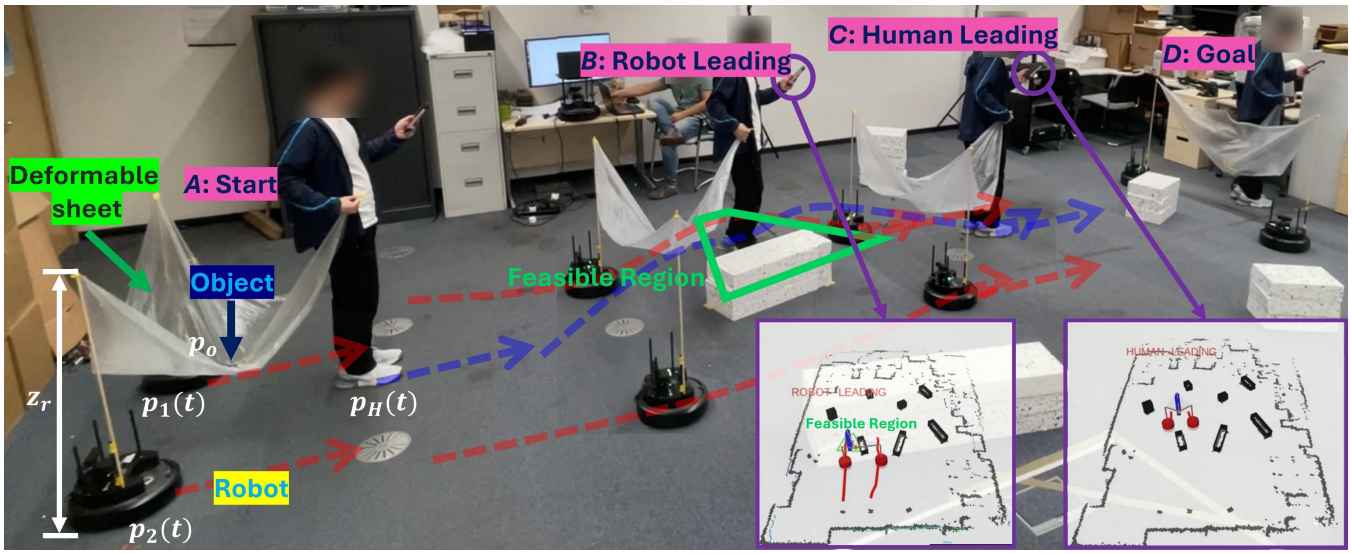

Fig. 1. A real-world demonstration of our human–robot collaborative transportation system. Here, a team of two nonholonomic robots and a human travels safely from point A to point D. The blue line is the human's predicted trajectory, and the red lines are the robot trajectories. At point B the formation operates in robot-leading mode to cross the obstacle. This constrains the human to a feasible region (shown in green) to avoid collisions. At point C the formation reverts to human-leading mode, where the robots follow the human.

***Abstract*— In this paper, we tackle real-time formation trajectory planning for collaborative object transportation in complex environments using a team of nonholonomic robots and a human. The object is transported in a *deformable sheet*, and robots should follow the human's lead while autonomously avoiding obstacles. By including a human in the formation, we leverage their adaptability and decision-making to improve transportation. However, it can be difficult for a human to predict how autonomous robots will behave in complex situations, such as when the formation must cross an obstacle, i.e. where the object is transported above it. This could cause human decisions that compromise safety. To overcome these challenges, we introduce a multi-modal formation planning framework. By default the human leads the formation, and the robots plan to remain in the same *homotopy class* as the human to avoid collisions. If obstacle crossing is necessary the robots take the lead of the formation, where human motion is constrained to a feasible region projected visually in front of them. We demonstrate the efficacy of our framework in simulation and on hardware.**

I. Introduction

Autonomous multi-robot formations have been widely used for collaborative object transportation in industrial manufacturing domains [1]. The safety and efficiency of transportation can be improved by adding a human who leads the transportation process alongside the robots. This allows us to leverage the human's adaptability and decision-making capabilities [2]. However, adding a human introduces new challenges. For example, the robots must continuously predict human motion intentions and proactively adapt their trajectories to assist the human and maintain safety [3]. There may also be situations where it is difficult for the human to lead the formation, such as when an obstacle should be crossed by rising the object above it (see Fig. 1). This maneuver requires the human to accurately predict the behavior of multiple robots. Therefore, leadership of the formation should change between the human and robots dynamically to maximize safety and improve transportation efficiency [4]. Crucially, the formation should be tolerant of failures caused by delayed robot responses or inaccuracies in human motion prediction [5]. This could be achieved through redundant manipulator joints [6]–[9] or specially designed mechanical structures [5], [10].

In this paper, we address collaborative object transportation for teams of nonholonomic robots and a human in cluttered environments. The object is transported in the center of a deformable sheet, where the human and robots grasp the corners of the sheet (see Fig. 1). Using a deformable sheet improves fault tolerance and flexibility, especially when handling irregular or fragile objects that cannot be directly grasped [11]. The sheet's height can also be adjusted by

Weijian Zhang, Charlie Street and Masoumeh Mansouri are with the School of Computer Science at the University of Birmingham, wxz163@student.bham.ac.uk and {c.l.street, m.mansouri}@bham.ac.uk
Charlie Street and Masoumeh Mansouri are UK participants in EU Project CONVINCE, and supported by UKRI grant number 10042096.

increasing the inter-agent distance, allowing the object to be carried above obstacles where necessary [12]. To solve formation motion planning in cluttered environments we propose a multi-modal real-time formation trajectory planning framework. By default, the human leads the formation (see points *A* and *C* in Fig. 1). Here, the robots predict the human's trajectory and optimize theirs to stay in the same homotopy class as the human, i.e. on the same side of obstacles, while maintaining safety. When obstacle crossing maneuvers are inevitable, i.e. where the object is transported above an obstacle, we transfer leadership over to the robots (see point *B* in Fig. 1), as it is challenging for the human to reason over robot behavior to guarantee object safety. While the robots are leading, the human is constrained to a safe feasible region that is visualized on a phone or through external projection (see the green region shown in Fig. 1). For safety, we also introduce a mode which realigns the formation if we are close to breaking it, e.g. if the human is about to move beyond the reach of the sheet.

Though human-robot collaborative transportation has been widely studied, ours is the first to consider a deformable sheet as the transportation mechanism. The primary contributions of this paper are as follows: (1) A multi-modal real-time formation trajectory planning framework where leadership switches between the human and robots to maximize safety and transportation efficiency; (2) A method for computing feasible regions that can be visualized to the human to guide them during obstacle crossing. We validate the efficacy of our framework in simulation and through real robot experiments. All corresponding software is released open source online[1].

II. Related Work

Human–Robot Formation Transportation. Collaborative object transportation is a significant topic in human-robot interaction. Here, we focus specifically on transportation using human-robot formations. For a comprehensive review of human-robot transportation methods, we refer the reader to [2]. For example, [8], [9] propose an adaptive, deformation-agnostic framework for human-robot transportation. There, compliant motions are generated for mobile manipulators assisting a human by combining human kinetic information with haptic signals sent from the object. However, these works assume obstacle-free environments. [7] presents a real-time motion planning framework for human-robot transportation which builds a capability map that describes feasible grasps for the mobile manipulator. Under this framework, the robot is always passive, adapting reactively to the human's motion. As a result, in extreme cases poor guidance from the human may cause the robot to fail. In [10], a bespoke mechanical sliding mechanism is attached to a mobile robot to simulate the kinematic redundancy of a human arm and improve fault tolerance. During transportation, the human guides the system toward the goal while the robot ensures obstacle avoidance along the human's trajectory. This approach is extended in [5] to switch the leader of the formation based on environmental sensory information. However, both [10] and [5] assume that obstacles are lower than the human hand, allowing the object to skim over them, limiting their applicability In contrast, our framework can be applied to dense, unstructured environments, where the object's height can be increased to clear high obstacles. Moreover, our framework is multi-modal, operating under human or robot leadership dependent on the situation. This maximizes safety and transportation efficiency.

Object Transportation using a Deformable Sheet. [13], [14] present formation planners for transporting deformable objects such as sheets, where the geometric constraints of the object are simplified to reduce complexity. In [15], a centralized planner is introduced for formations of car-like robots transporting objects in a net in unstructured environments. However, [13]–[15] assume the formation is a rigid 2D body that must remain strictly separated from obstacles to avoid collisions. This assumption can severely limit the solution space in cluttered environments [16]. To facilitate more flexible obstacle avoidance, several works allow formations to *cross* obstacles by letting the robots navigate on different sides of it while the object rides above it. For example, [17]–[19] use three omnidirectional robots to grasp and support a deformable sheet for object transportation. Here, the height of the object in the sheet increases with the inter-robot distance. In [11], the interaction model between the sheet and object is approximated with a set of virtual, variable-length cables. This is different to conventional cable-driven approaches that assume fixed-length tethers [20], [21]. A local planner is then proposed to facilitate obstacle crossing. In [16], a formation-level global path planner is presented for navigation in cluttered environments, though its success rate heavily depends on finely the environment is discretized. [11], [16]–[19] all rely on holonomic robots. For formations of nonholonomic robots, [12] introduces a global trajectory planner based on the forward kinematics model in [22], enabling collaborative transportation and obstacle crossing in cluttered environments. Compared to the above works, our framework explicitly accounts for nonholonomic robots and supports obstacle crossing as well as detouring around them. Our approach also functions as a local planner for online trajectory synthesis, which admits real-time human-robot collaborative transportation.

III. Problem Statement

In this paper, we consider a formation of two differential-drive nonholonomic robots and a human transporting an object using a deformable sheet. However, our framework is general and can be applied to robots with arbitrary kinematics, including holonomic and nonholonomic platforms. Let $\mathcal{W} \subset \mathbb{R}^3$ be the $3D$ workspace, and $\mathcal{O} \subset \mathbb{R}^3$ be the set of static convex polyhedral obstacles which are known *a priori*. During object transportation, the robots move in $2D$ space, and the transported object moves in $3D$. To simplify object transportation, we apply the assumptions in [11], i.e. the object is a point mass, the object's motion is quasi-static, and the deformable sheet is inelastic, such that the

[1] https://github.com/HyPAIR/HRCTF

object moves freely under gravity towards the position of minimum potential energy on the sheet. For robot $i \in \{1, 2\}$, $\mathbf{p}_i(t) = (x_i(t), y_i(t))^T \subset \mathbb{R}^2$ is the coordinate of the midpoint of the rear axle and $\theta_i(t) \in [-\pi, \pi)$ is the robot's orientation in Cartesian space. The human state at time $t \in [t_0, t_0 + t_f]$ is $\mathbf{z}_H(t) = [\mathbf{p}_H(t), \theta_H(t), \mathbf{v}_H(t)]$, where $\mathbf{p}_H(t) = [x_H(t), y_H(t)]^T$ denotes the position, $\theta_H(t)$ the orientation, and $\mathbf{v}_H(t)$ the velocity. Here, t_0 is the start of the prediction horizon, and t_f is the horizon length. With this, the human's predicted trajectory is given by $\mathcal{T}_H = \{\mathbf{p}_H(t)\}_{t=t_0}^{t_0+t_f}$. The object's position is denoted $\mathbf{p}_o \subset \mathbb{R}^3$. A $3D$ formation view is shown in Fig. 1. We capture the formation topology by labeling the agents as $\{1, 2, H\}$ in an anticlockwise direction. This ordering must remain constant during transportation to avoid the sheet folding or twisting. We also define a template for the formation that captures its desired shape. This template is given as a set of positions for each agent $\mathbf{p}^{\text{form}}_{\{1,2,H\}}$ relative to the formation center. Each robot's support point to the sheet is fixed at a constant height z_r. We adopt a unicycle model [23] to describe robot kinematics, assuming pure rolling without lateral slipping. A trajectory for robot i is given by $\mathcal{T}_i(t) = \{[\mathbf{z}_i(t), \mathbf{u}_i(t)]^T\}_{t=t_0}^{t_0+t_f}$, where $\mathbf{z}_i(t) = [\mathbf{p}_i(t), \theta_i(t)]$ and $\mathbf{u}_i(t)$ are the state and control parameters.

The object transportation problem tasks us with finding trajectories $\mathcal{T}_i$ for each robot such that the object reaches target configuration $\mathbf{z}^g$ from initial configuration $\mathbf{z}^s$. Moreover, we must explicitly assist and guide the human to ensure the safety of the object and the sheet. We formulate this as a joint optimal control problem (OCP):

$$\min_{\mathbf{z}_i, \mathbf{u}_i} \sum_i \int_{t_0}^{t_0+t_f} \mathbf{u}_i(t)^T \mathbf{W} \mathbf{u}_i(t)\, dt \tag{1a}$$

$$\text{s.t. kinematic constraints,} \tag{1b}$$

$$\underline{\mathbf{z}} \le \mathbf{z}_i(t) \le \overline{\mathbf{z}},\ \underline{\mathbf{u}} \le \mathbf{u}_i(t) \le \overline{\mathbf{u}}, \tag{1c}$$

$$\mathbf{z}_i(t_0) = \mathbf{z}_i^s,\ \mathbf{u}_i(t_0) = \mathbf{u}_i^s, \tag{1d}$$

$$\mathbf{z}_i(t_0 + t_f) = \mathbf{z}_i^g,\ \mathbf{u}_i(t_0 + t_f) = \mathbf{u}_i^g, \tag{1e}$$

$$\mathcal{G}(\mathbf{z}_i(t), \mathbf{u}_i(t), t) \preceq 0, \tag{1f}$$

$$\forall t \in [t_0, t_0 + t_f], \quad \forall i \in \{1, 2\},$$

where $\mathbf{W} \in \mathbb{R}^{2\times 2}$ is a diagonal matrix that penalizes control effort; $\underline{\mathbf{z}}$, $\overline{\mathbf{z}}$, $\underline{\mathbf{u}}$, and $\overline{\mathbf{u}}$ are lower and upper bounds on the states and control inputs, respectively; $\mathbf{z}_i^s$ and $\mathbf{z}_i^g$ are the start and goal states for robot i; and $\mathbf{u}_i^s$ and $\mathbf{u}_i^g$ are the control inputs at robot i's start and goal states, respectively. The constraints enforce robot kinematics (1b), state and input bounds (1c), and boundary conditions (1d)–(1e) (1f) comprises robot safety constraints, including obstacle-avoidance constraints, inter-robot collision-avoidance constraints [24], size and topological safety constraints [15], and object height [22] and terminal constraints. For simplicity, in this paper we only consider formations with two robots and a human. However, our approach is general and can be applied to larger teams of robots. We will investigate the scalability of our framework on larger teams in future work.

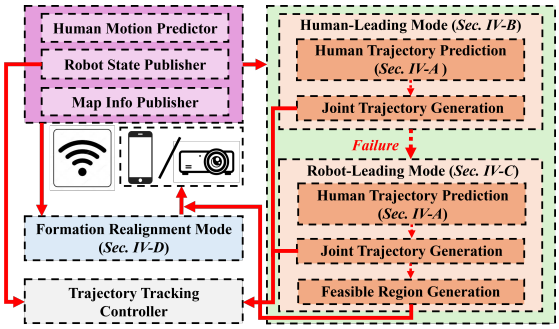

Fig. 2. The proposed collaborative transportation framework.

IV. Human-Robot Transportation Framework

In Fig. 2, we present our online receding-horizon framework for collaborative human-robot object transportation using a deformable sheet. Here, a centralized node periodically broadcasts the human's predicted motion, robot states, and global map information. Upon receiving this information, the trajectory planner predicts the human's future trajectory (Sec. IV-A) and attempts to generate a formation trajectory in human-leading mode (Sec. IV-B). If a feasible solution cannot be obtained within the current horizon, the framework switches to robot-leading mode (Sec. IV-C), where a feasible region is computed and projected in real time to guide the human's motion. The resulting trajectory is sent to the trajectory tracking controller for execution. If an infeasible formation is detected due to unexpected events such as sudden human movement or disturbances, the system switches to formation realignment mode (Sec. IV-D), which restores the desired configuration before resuming the receding-horizon planning process.

A. Human Motion Prediction

We predict the human's trajectory for each planning horizon based on the human's current state and the surrounding environment. For this, we use the intention-aware hybrid A^* approach in [25]. For each horizon, the human's local goal is computed by simulating its motion from its current position, i.e., $\mathbf{p}_H(t_0 + t_f) = \mathbf{p}_H(t_0) + t_f \cdot \mathbf{v}_H(t_0)$. If $\mathbf{p}_H(t_0 + t_f)$ lies inside an obstacle, we extend the horizon by δt until a collision-free position is identified. The hybrid A* then searches over a set of human motion primitives, where each primitive corresponds to an intention for the human. The final path aims to minimise intention switching by the human. The corresponding predicted trajectory is then stored in $\mathcal{T}_H$.

B. Human-Leading Mode

We summarize the operation of human-leading mode in Alg. 1. In this mode, the formation must traverse around obstacles rather than cross them. This is because it is challenging for the human to reliably infer feasible motions for the formation that guarantees collision avoidance for the object. Given the human's predicted trajectory as described in Sec. IV-A (lines 1-2 in Alg. 1), we construct a safe corridor along the human's predicted trajectory $\mathcal{T}_H$ using the method in [26]. This ensures each waypoint along $\mathcal{T}_H$ is enclosed within a convex, obstacle-free region. The method in [26]

expands collision-free line segments in $\mathcal{T}_H$ into a series of connected convex polygons which can be represented as [27]:

$$\mathcal{SC}^H = \{q \in \mathbb{R}^2 \mid \mathbf{A}^H \cdot q \leq \mathbf{b}^H\}, \tag{2}$$

where $\mathbf{A}^H \in \mathbb{R}^{n\times 2}$ and $\mathbf{b}^H \in \mathbb{R}^n$ define the hyperplanes of the polygon, and n denotes the number of hyperplanes. The union of these polygons define the human's safe corridor $\mathcal{SC}^H$ (line 3 in Alg. 1), shown as the light-blue regions in Fig. 3. By constraining robot trajectories to $\mathcal{SC}^H$, the robots remain in the same homotopy class as the human, which guarantees safe traversal around obstacles.

Given the safe corridor, we now compute a local goal for each robot that lies within $\mathcal{SC}^H(t_0 + t_f)$. The local goal for robot i can be expressed as $s \cdot \mathbf{p}_i^{\text{form}} + \mathbf{d}$, i.e. in terms of its relative position in the formation template. Here, $s \in \mathbb{R}$ is a scaling factor that expands or contracts the formation, and $\mathbf{d} \in \mathbb{R}^2$ is a translation vector. s is set to 1 by default. If $s = 1$ and $\mathbf{d}$ is the vector between the formation centers at the start and end of the horizon, $s\cdot\mathbf{p}_i^{\text{form}}+\mathbf{d}$ perfectly positions robot i with respect to the human and formation template at time t_0+t_f. However, this position may be blocked, and so we solve a convex quadratic program to optimize s and $\mathbf{d}$ (line 4):

$$\min_{s,\mathbf{d}} (s-1)^2 \tag{3a}$$

$$\text{s.t. } \mathbf{A}^H(t_0+t_f)\cdot(s\cdot\mathbf{p}_i^{\text{form}}+\mathbf{d}) \leq \mathbf{b}^H(t_0+t_f), \tag{3b}$$

$$\mathbf{p}_H(t_0+t_f) = s\cdot\mathbf{p}_H^{\text{form}}+\mathbf{d}, \tag{3c}$$

$$\underline{\mathbf{s}} \leq \mathbf{s} \leq \bar{\mathbf{s}}, \tag{3d}$$

where objective (3a) penalizes deviations from $s = 1$ to keep the formation close to its desired shape. Constraint (3b) constrains the local goals for each robot to the safe corridor, and (3c) fixes the human's local goal.

After synthesizing the local goal for each robot, we generate a corresponding joint trajectory which acts as the initial guess for OCP (1) (line 6). For this, we use an adapted version of hybrid A^* [28] where child nodes are only expanded if they remain in the safe corridor. This leads to collision-free paths that satisfy robot kinodynamic constraints while remaining homotopically equivalent with the human. These paths are then augmented with a minimum-time velocity profile [29] to produce trajectories, where the robot velocities are proportionally relaxed such that they reach their targets at the same time as the human [30]. Finally, we solve (1) in line 9 to synthesize the final robot trajectories. If we fail to synthesize local goals in (3) or trajectories in (1), there is likely no feasible solution in the safe corridor. This may occur in narrow environments. To find a solution, we must consider alternate homotopy classes for the robots, necessitating obstacle crossing. Therefore, when this occurs we switch to robot-leading mode (lines 5 and 11), which we discuss next.

C. Robot-Leading Mode

In robot-leading mode, the human is guided under robot leadership to successfully cross an obstacle. The operation of robot-leading mode is summarized in Alg. 2.

Algorithm 1: Human-Leading Mode.

1 $\mathbf{p}_H(t_0+t_f) \leftarrow$ **ForwardSimulation**$(\mathbf{p}_H(t_0))$;
2 $\mathcal{T}_H \leftarrow$ **PredictHumanTraj**$(\mathbf{p}_H(t_0), \mathbf{p}_H(t_0+t_f))$;
3 $\mathcal{SC}^H \leftarrow$ **GenerateCorridors**$(\mathcal{T}_H)$;
4 **if** *No solution to* (3) **then**
5 | *Switch to robot-leading mode;*
6 $\forall i, \mathcal{T}_i^g \leftarrow$ **InitGuess**$(\mathbf{p}_i(t_0), \mathbf{p}_i(t_0+t_f), \mathcal{SC}^H)$;
7 **GenerateOCP**$(\mathcal{T}_1^g, \mathcal{T}_2^g, \mathcal{SC}^H)$;
8 **if** *Solution to* (1) *found* **then**
9 | **return** $\mathcal{T}_1, \mathcal{T}_2$;
10 **else**
11 | *Switch to robot-leading mode;*

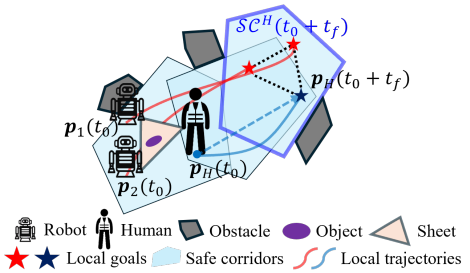

Fig. 3. Constructing trajectories in human-leading mode.

1) Iterative Trajectory Optimization: Similar to human-leading mode, we begin by computing a local goal for each robot. For this, we simulate the human and robots moving in a straight line under the human's velocity for the horizon duration to obtain temporary local goals $\mathbf{p}_{\{i,H\}}^{\text{temp}}$ (see Fig. 4(a) and line 1 of Alg. 2). If the temporary local goals are blocked, we incrementally extend the horizon by δt, as in Sec. IV-A. Similar to Sec. IV-B, we then obtain the local goals by constructing the safe corridor for each robot (line 2), and optimizing the scaling factor s and translation vector $\mathbf{d}$ through the following convex quadratic program (line 3):

$$\min_{s,\mathbf{d}} \text{ (3a)} \tag{4a}$$

$$\text{s.t. } \mathbf{A}^i(t_0+t_f)\cdot(s\cdot\mathbf{p}_i^{\text{form}}+\mathbf{d}) \leq \mathbf{b}^i(t_0+t_f), \tag{4b}$$

$$\text{(3c), (3d).} \tag{4c}$$

The only difference between (3) and (4) is that constraint (4b) constrains each robot's local goal to their own safe corridor rather than the human's. After synthesizing the local goals we predict the human's trajectory as in Sec. IV-A (line 4), and then explore different homotopy classes for the robots to increase our chance of finding a feasible joint trajectory. For this, we construct a topological probabilistic roadmap (PRM) [31] using the centers of the initial and goal formations(line 5 and Fig. 4(b)). From the topological PRM, we can extract multiple homotopically distinct paths for each robot (line 6) and then combine them to give candidate joint paths (line 6) for trajectory synthesis. However, the number of homotopy classes increases exponentially with the number of obstacles, making it computationally challenging to enumerate through all candidate joint paths. Therefore, we filter and rank the candidate joint paths (line 7). We discretize

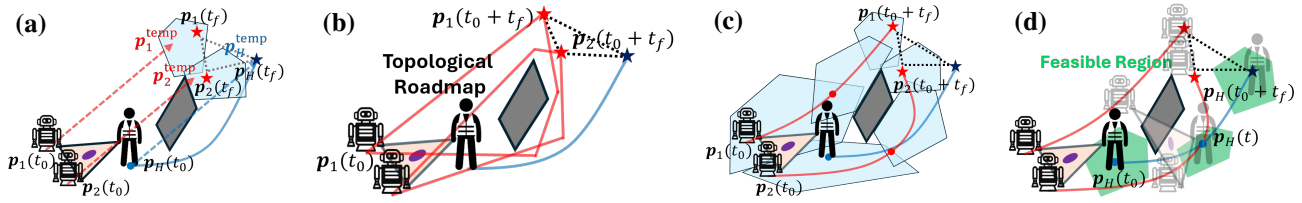

Fig. 4. Trajectory synthesis in robot-leading mode. (a) Temporary local goal generation using forward simulation at a constant velocity. The red lines show the simulated trajectories, and the light blue polygons are the safe corridors at each robot's temporary local goal. (b) Initial guess generation. Red paths are homotopically distinct paths for each robot. (c) Trajectory synthesis. Robot trajectories are in red, and the human trajectory is in blue. The blue polygons are safe corridors generated from the initial guess. (d) Feasible region generation, where the projected regions are shown in green.

Algorithm 2: Robot-Leading Mode.

1 $\mathbf{p}^{\text{temp}}_{\{i,H\}} \leftarrow$ **ForwardSimulation**$(\mathbf{p}_{\{i,H\}}(t_0))$;
2 $\forall i,\ \mathcal{SC}^i(t_0+t_f) \leftarrow$ **GenerateCorridors**$(\mathbf{p}^{\text{temp}}_i)$;
3 $\mathbf{p}_{\{i,H\}}(t_0+t_f) \leftarrow$ **Solve** (4);
4 $\mathcal{T}_H \leftarrow$ **PredictHumanTraj**$(\mathbf{p}_H(t_0),\ \mathbf{p}_H(t_0+t_f))$;
5 $\forall i,\ \mathbf{path}_i \leftarrow$ **TopoPRM**$(\mathbf{p}_i(t_0),\ \mathbf{p}_i(t_0+t_f))$;
6 $\mathbf{comb} \leftarrow$ **IdentifyComb**$(\mathbf{path}_1, \mathbf{path}_2)$;
7 $\mathbf{comb}^* \leftarrow$ **Filtering&Sorting**$(\mathbf{comb})$;
8 **for** $\mathbf{comb_j} \in \mathbf{comb}^*$ **do**
9 $\quad \mathcal{SC}^1, \mathcal{SC}^2 \leftarrow$ **GenerateCorridors**$(\mathbf{comb_j})$;
10 $\quad \mathcal{T}_1^g, \mathcal{T}_2^g \leftarrow$ **GenerateInitGuess**$(\mathbf{comb_j}, \mathcal{SC}^1, \mathcal{SC}^2)$;
11 $\quad$ **GenerateOCP**$(\mathcal{T}_1^g, \mathcal{T}_2^g, \mathcal{SC}^1, \mathcal{SC}^2)$;
12 $\quad \mathcal{T}_1, \mathcal{T}_2 \leftarrow$ **Solve** (1);
13 $\quad$ **while** *Height constraints not satisfied* **do**
14 $\quad\quad$ *Increase minimum inter-robot distance* d_{min};
15 $\quad\quad$ **if** $d_{min} >$ *maximum inter-robot distance* **then**
16 $\quad\quad\quad$ **return** $\emptyset$;
17 $\quad\quad \mathcal{T}_1, \mathcal{T}_2 \leftarrow$ **Solve** (1);
18 $\quad$ **return** $\mathcal{T}_1, \mathcal{T}_2$;

joint paths into waypoints aligned with the human trajectory, and discard candidates if they cross obstacles that are too high to be crossed by the sheet, or if the inter-robot distances exceed the sheet's maximum range. The remaining candidate joint paths are then sorted by total path length, and iterated through until we find a feasible joint trajectory. For each candidate joint path, we build safe corridors for each robot that constrain them to the homotopy class of that path (line 9). We then generate an initial guess joint trajectory within these safe corridors in line 10 using the same hybrid A^* approach used for human-leading mode (see Sec. IV-B). Given the initial guess, we then try to synthesize a joint trajectory. Computing the object height requires solving a forward kinematics function that returns the object's equilibrium state in the sheet given the states of the robots [22]. This forward kinematics function approximates the sheet using a virtual variable cable model [11]. Adding this directly into OCP (1) would significantly increase complexity. To mitigate this, we express height constraints in terms of the inter-agent distance, where the height of the object increases as the agents move apart. Given the complexity of mapping from the inter-agent distance to object height we solve OCP (1) iteratively (lines 13-18). This is because additional optimization variables and

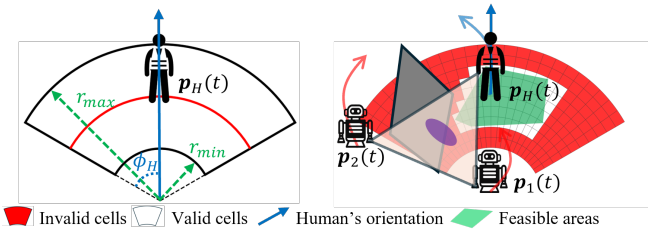

Fig. 5. Generating the projected feasible region for the human.

complex nonlinear constraints are introduced which make OCP (1) difficult to solve directly. For further details, we refer the reader to [12]. Given an initial constraint on the minimum inter-robot distance, we try to solve OCP (1) (line 12). Upon finding a solution, we validate whether the object height constraints are satisfied using the forward kinematics function in [22] (line 13). If the constraints are satisfied, we return the joint trajectory (line 18). Otherwise, we incrementally increase the minimum inter-robot distance (line 14), increasing the object height, and try to solve OCP (1) again (line 17). We repeat this until a solution is found or a maximum inter-robot distance is reached. An example joint trajectory is shown alongside the safe corridors in Fig. 4(c).

2) Joint Trajectory Freezing: A single obstacle crossing maneuver may occur over multiple planning horizons. To reduce unnecessary human movement and confusion, we want to keep each agent in the same homotopy class until the obstacle is cleared. To guarantee this, after finding a solution with Alg. 2, we freeze the trajectory such that it executes until completion, or until Alg. 1 can find a solution that enables a smooth transition back to human-leading mode. This is demonstrated in Fig. 1, where the formation switches to robot-leading mode at point B, executes the frozen trajectory, and then switches back to human-leading mode at point C, as Alg. 1 has found a feasible solution.

3) Feasible Region Generation: While executing obstacle crossing maneuvers, it may be challenging for the human to identify safe motions that satisfy the object height constraints. To aid with this, we project a feasible region for the human at each timestep via a phone or external projector. If the human keeps within this region, system safety is guaranteed. Inspired by [32], we define the feasible region as a 2D annular region centered at the human's current position $\mathbf{p}_H(t)$. An annular region is a segment of a ring around $\mathbf{p}_H(t)$ with an inner radius of $r_{\min}$, an outer radius of $r_{\max}$,

Algorithm 3: Formation Realignment Mode.

1 **if** $f_s(\mathbf{p}_i(t),\, \mathbf{p}_H(t)) > \epsilon_{sim}$ **then**
2 $\quad \mathcal{SC}^H(t) \leftarrow$ **GenerateCorridors**$(\mathbf{p}_H(t))$;
3 $\quad \mathbf{p}_H^{re}(t) \leftarrow$ **Solve** (6);
4 $\quad$ **while** *Human distance to* $\mathbf{p}_{re}(t) > e_{dis}$ **do**
5 $\quad\quad$ *Robots remain stationary*;
6 $\quad$ *Switch to other modes*;

and a neighborhood angle of ϕ_H, which defines the angular range to the left and the right of the human (see the left of Fig. 5). We discretize this annular region into cells and mark cells as infeasible if they are occupied or would violate formation, topology, or height constraints (see the red cells on the right of Fig. 5). We then treat the infeasible cells as obstacles, and apply the method in [26] to compute a polygonal convex feasible region using the human's current position as the seed point. This region is then projected to the human, and is updated at each timestep, as shown in Fig. 4(d).

D. Formation Realignment Mode

During transportation, unexpected events may occur that the robots cannot respond to quickly enough, compromising the safety of the system. For example, the human may make a sudden unexpected movement, or the human prediction system may produce large errors. To address this, we introduce a third mode to our framework for formation realignment, described in Alg. 3 and Fig. 6. We switch to this mode when the current formation configuration deviates too far from the formation template introduced in Sec. III (line 1 in Alg. 3). To quantify this deviation, we adopt the differentiable similarity error metric in [33]:

$$f_s(\mathbf{p}_i(t),\, \mathbf{p}_H(t)) = \|\hat{\mathbf{L}} - \hat{\mathbf{L}}_{\text{form}}\|_F^2, \tag{5}$$

where $\hat{\mathbf{L}}$ and $\hat{\mathbf{L}}_{\text{form}}$ denote the symmetric normalized Laplacian matrices of the current and desired formation configurations, and $|\cdot|_F$ denotes the Frobenius norm. This metric captures the similarity of shape between the two formation configurations independent of scale, translation, and rotation. Once we enter formation realignment mode, the robots stop until the human reaches a safe position to continue (lines 4-5). Similar to robot-leading mode, this safe position is projected to the human. To compute the safe position, we compute the safe corridor $\mathcal{SC}^H$ around the human's current position (line 2), and then solve a constrained nonlinear program (NLP) (line 3):

$$\min_{\mathbf{p}_H^{re}(t)} \; f_s(\mathbf{p}_i(t),\, \mathbf{p}_H^{re}(t)), \tag{6a}$$

$$\text{s.t. } \mathbf{A}^H(t) \cdot \mathbf{p}_H^{re}(t) \leq \mathbf{b}^H(t). \tag{6b}$$

Here, the objective (6a) minimizes formation deviation and constraint (6b) keeps the human inside the safe corridor, guaranteeing collision avoidance. Once the human approaches the safe position, we switch back to human-leading or robot-leading mode (line 6).

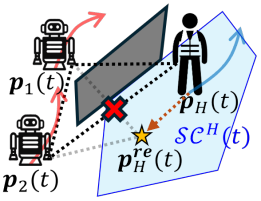

Fig. 6. Formation realignment mode. The formation is about to fail as the human is moving too far from robot 2 and approaching the maximum range of the sheet. The safe human position (the orange star) is obtained by solving (6).

V. Experimental Results

In this section, we demonstrate the efficacy of our approach in simulation and on hardware. All simulations are run on Ubuntu 20.04 with an Intel Core i9 processor @ 3.2GHz and 16GB of RAM. All software is implemented in C++ using the robot operating system (ROS). We use a unified pipeline to solve the optimization problems in (3), (4), (6), and the OCP in (1). The OCP is discretized with the explicit Runge–Kutta method, whereas (3), (4), and (6) are formulated directly. All problems are differentiated with ADOL-C [34] and solved by IPOPT [35]. Experimental runs can be viewed in the supplementary video. All parameters are hand tuned to maximize performance. Parameter values are provided in our software documentation.

A. Simulation Experiment

We begin by analysing framework performance in Gazebo, where a joystick is used to emulate human motion. Here, the maximum human velocity and angular velocity are set to $0.8\,m/s$ and $0.5\,rad/s$, respectively. We consider three $60 \times 60\,m^2$ environment classes with fixed initial and target formation configurations. The first class has 50 obstacles (sparse), the second has 70 (medium), and the third has 90 (dense). We randomly generate 35 environments for each class, where each random environment corresponds to a single problem instance. For each instance, we evaluate the success rate (SR), average task completion time (CT), average robot trajectory length (TL), and the average/maximal formation similarity error $\bar{e}_{sim}/\hat{e}_{sim}$ as in (5), which describes how much the formation adjusts its shape. We present our results in Table I, where the values for CT, TL, $\bar{e}_{sim}$, and $\hat{e}_{sim}$ only reflect successful runs with no safety constraint violations. Our framework achieves a very high success rate across all environments. The task completion time, trajectory length, and formation similarity error rise with the obstacle density, as more frequent detours and obstacle crossing is required for collision avoidance. When failures occurred, these were often caused by the human operator inadvertently guiding the formation into very narrow areas where no feasible solution could be found.

To the best of our knowledge, no prior work has addressed collaborative transportation using a human–robot team with a deformable sheet. Existing work on human-robot transportation either relies on mobile manipulators [7]–[9] or bespoke mechanical extensions to the robots [5], [10]. Transportation approaches that use a deformable sheet have only considered

TABLE I

PERFORMANCE AS THE NUMBER OF OBSTACLES INCREASES.

Env.	SR(%)	CT (s)	TL (m)	$\bar{e}_{sim}$	$\hat{e}_{sim}$
Sparse	100	85.073	72.914	0.019	0.072
Medium	97.1	88.984	75.297	0.032	0.106
Dense	91.4	92.914	79.213	0.045	0.108

TABLE II

ABLATION STUDY.

Method.	SR(%)	CT (s)	TL (m)	$\bar{e}_{sim}$	$\hat{e}_{sim}$
w\o **RL**	68.8	101.275	87.553	**0.023**	**0.049**
w\o **TF**	81.3	97.312	81.358	0.054	0.142
w\o **FR**	43.7	95.246	82.280	0.041	0.086
Proposed	**100**	**92.914**	**79.213**	0.045	0.108

entirely robot-based teams [11]–[19], [22]. With this, there are no approaches we can directly compare against, and so we conduct an ablation study to evaluate the efficacy of each component of our framework. We consider the following variants of our framework: **w/o robot leading mode (RL)** - We only use human-leading mode, prohibiting obstacle crossing; **w/o trajectory freezing (TF)** - All modes are used, but in robot-leading mode trajectories are re-planned at each horizon rather than frozen; **w/o feasible region projection (FR)** - All modes are used, but no feasible region is projected to the human in robot-leading mode.

We test each of these variations on the 32 successful runs for dense environments in the previous experiment. The average computation time per planning horizon for each run is under 250ms, admitting real-time use of our framework. We present our results in Table II. When we remove robot-leading mode (Sec. IV-C) we struggle to find feasible solutions in dense environments, leading to task failure. Further, as obstacles cannot be crossed, the formation makes frequent detours, increasing the completion times and trajectory lengths. Though obstacle crossing sacrifices the formation similarity, it increases formation flexibility, improving success rate, completion time, and trajectory length. When we remove trajectory freezing (Sec. IV-C.2), the formation struggles to maintain topological consistency during obstacle crossing. Frequent switches in homotopy class cause unnecessary maneuvers, increasing the completion time, trajectory length, and formation similarlity error. When we stop projecting feasible regions during obstacle crossing (Sec. IV-C.3), the human often fails to follow a safe trajectory, violaing safety constraints and causing task failure. In contrast, our framework consistently completes transportation with the shortest completion time and trajectory length. We visualize some example results from the ablation study in Fig. 7.

B. Real-World Experiment

In addition to simulated experiments, we validated our framework on hardware in an indoor environment with cuboid obstacles (see Fig. 1 and Fig. 8). The experiments were conducted using two TurtleBot 4 differential-drive robots. Due to hardware limitations, we visualize the feasible region to the human using a phone rather than a projection on the floor. We use DATMO [36] on an external 2D LiDAR to track the human and estimate its position, orientation, and velocity. Robot states are obtained through ROS. Our framework updates and publishes trajectories every one second (the planning horizon), which are then tracked by a closed-loop controller running at 100Hz. We consider five real-world scenarios. In all experiments, the human maintains the sheet at a fairly constant height, where small deviations are handled safely via a conservative height margin in the solver constraints. In Fig. 1, we show an example of obstacle crossing, which requires switching to and from robot-leading mode. In Fig, 8(a), we consider a scenario where the formation traverses two obstacles consecutively, staying in robot-leading mode throughout. In Fig. 8(b), the formation adjusts its shape to safely pass through a narrow corridor before crossing an obstacle. Finally, in Fig. 8(c), the human makes a sudden movement forward which triggers formation realignment mode. Once the human reaches a safe position, the system switches back to human-leading mode and completes the task. Note that while Fig. 8(c) shows the human stepping forward, the same realignment process is triggered by unexpected motion in any direction. These results demonstrate our framework's ability to adapt naturally to different obstacle landscapes and environments. Our experiments are shown in full in the supplementary video.

VI. CONCLUSION

In this paper, we presented a formation planning framework for collaborative object transportation using two nonholonomic robots and a human with a deformable sheet. In future work, we will scale to larger formations capable of transporting heavier objects. For this, we will explore how to enable collaboration in multi-human multi-robot formations, and how to model the dynamics of heavy objects.

REFERENCES

[1] E. Tuci, M. H. Alkilabi, and O. Akanyeti, "Cooperative object transport in multi-robot systems: A review of the state-of-the-art," *Frontiers in Robotics and AI*, vol. 5, p. 59, 2018.

[2] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, "Progress and prospects of the human–robot collaboration," *Autonomous robots*, vol. 42, no. 5, pp. 957–975, 2018.

[3] J. E. Domínguez-Vidal and A. Sanfeliu, "Improving human-robot interaction effectiveness in human-robot collaborative object transportation using force prediction," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 7839–7845, IEEE, 2023.

[4] A. J. Mahmud, W. Li, and X. Wang, "Mutual adaptation in human-robot co-transportation with human preference uncertainty," *arXiv preprint arXiv:2503.08895*, 2025.

[5] N. Feng, X. Guo, X. Yu, S. Zhang, and W. He, "Human-robot sharing operation in co-transporting for nonholonomic mobile robot," *IEEE Transactions on Cognitive and Developmental Systems*, 2024.

[6] D. Sirintuna, T. Kastritsi, I. Ozdamar, J. M. Gandarias, and A. Ajoudani, "Enhancing human–robot collaborative transportation through obstacle-aware vibrotactile warning and virtual fixtures," *Robotics and Autonomous Systems*, vol. 178, p. 104725, 2024.

[7] H. Zhang, Q. Sheng, J. Hu, X. Sheng, Z. Xiong, and X. Zhu, "Cooperative transportation with mobile manipulator: A capability map-based framework for physical human–robot collaboration," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 6, pp. 4396–4405, 2022.

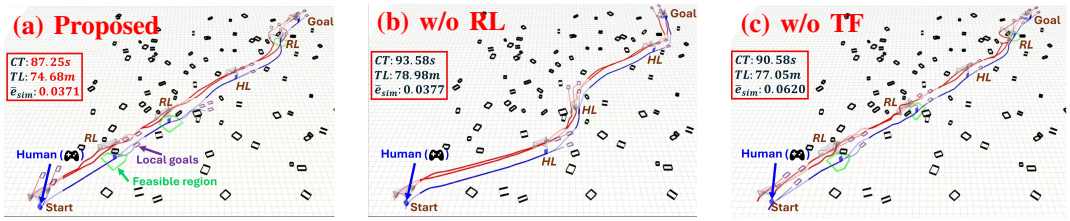

Fig. 7. Trajectories for each variation of our framework in the ablation study. "HL" and "RL" denote human-leading and robot-leading mode, respectively.

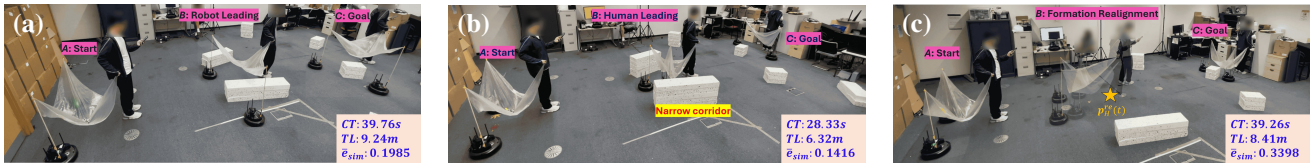

Fig. 8. Snapshots from three real-world experiments alongside the quantitative results. (a) Robot-leading mode only; (b) traversing a narrow corridor; (c) formation realignment mode.

[8] D. Sirintuna, I. Ozdamar, and A. Ajoudani, "Carrying the uncarriable: a deformation-agnostic and human-cooperative framework for unwieldy objects using multiple robots," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7497–7503, 2023.

[9] D. Sirintuna, A. Giammarino, and A. Ajoudani, "An object deformation-agnostic framework for human–robot collaborative transportation," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 2, pp. 1986–1999, 2023.

[10] X. Yu, X. Guo, W. He, M. A. Mughal, and D. Zhang, "Real-time trajectory planning and obstacle avoidance for human–robot co-transporting," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 2969–2985, 2024.

[11] J. Hu, W. Liu, H. Zhang, J. Yi, and Z. Xiong, "Multi-robot object transport motion planning with a deformable sheet," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9350–9357, 2022.

[12] W. Zhang, C. Street, and M. Mansouri, "Multi-nonholonomic robot object transportation with obstacle crossing using a deformable sheet," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 7349–7355, 2025.

[13] J. Alonso-Mora, R. Knepper, R. Siegwart, and D. Rus, "Local motion planning for collaborative multi-robot manipulation of deformable objects," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5495–5502, IEEE, 2015.

[14] S. Chand, P. Verma, R. Tallamraju, and K. Karlapalem, "Transportation of deformable payload through static and dynamic obstacles using loosely coupled nonholonomic robots," in *Proceedings of the International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*, pp. 228–230, IEEE, 2019.

[15] L. Pei, J. Lin, Z. Han, L. Quan, Y. Cao, C. Xu, and F. Gao, "Collaborative planning for catching and transporting objects in unstructured environments," *IEEE Robotics and Automation Letters*, 2023.

[16] W. Liu, J. Hu, H. Zhang, M. Y. Wang, and Z. Xiong, "A novel graph-based motion planner of multi-mobile robot systems with formation and obstacle constraints," *IEEE Transactions on Robotics*, 2023.

[17] K. Hunte and J. Yi, "Collaborative object manipulation through indirect control of a deformable sheet by a mobile robotic team," in *Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 1463–1468, IEEE, 2019.

[18] K. Hunte and J. Yi, "Collaborative manipulation of spherical-shape objects with a deformable sheet held by a mobile robotic team," *IFAC-PapersOnLine*, vol. 54, no. 20, pp. 437–442, 2021.

[19] K. Hunte and J. Yi, "Pose control of a spherical object held by deformable sheet with multiple robots," *IFAC-PapersOnLine*, vol. 55, no. 37, pp. 414–419, 2022.

[20] Y. Liu, Z. Cao, H. Xiong, J. Du, H. Cao, and L. Zhang, "Dynamic obstacle avoidance for cable-driven parallel robots with mobile bases via sim-to-real reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1683–1690, 2023.

[21] T. Rasheed, P. Long, A. S. Roos, and S. Caro, "Optimization based trajectory planning of mobile cable-driven parallel robots," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6788–6793, IEEE, 2019.

[22] J. Hu, W. Liu, J. Yi, and Z. Xiong, "Forward kinematics of object transporting by a multi-robot system with a deformable sheet," *IEEE Robotics and Automation Letters*, 2024.

[23] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to autonomous mobile robots*. MIT press, 2011.

[24] W. Zhang, C. Street, and M. Mansouri, "A decoupled solution to heterogeneous multi-formation planning and coordination for object transportation," *Robotics and Autonomous Systems*, p. 104773, 2024.

[25] Q. Ren, H. Yu, J. Dai, Z. Zheng, J. Meng, L. Xu, C. Xu, F. Gao, and Y. Cao, "Intention-aware planner for robust and safe aerial tracking," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 8153–8160, IEEE, 2024.

[26] S. Liu, M. Watterson, K. Mohta, K. Sun, S. Bhattacharya, C. J. Taylor, and V. Kumar, "Planning dynamically feasible trajectories for quadrotors using safe flight corridors in 3-d complex environments," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1688–1695, 2017.

[27] C. D. Toth, J. O'Rourke, and J. E. Goodman, *Handbook of discrete and computational geometry*. CRC press, 2017.

[28] D. Dolgov, S. Thrun, M. Montemerlo, and J. Diebel, "Path planning for autonomous vehicles in unknown semi-structured environments," *The Int. Journal of Robotics Research*, vol. 29, no. 5, pp. 485–501, 2010.

[29] T. Lipp and S. Boyd, "Minimum-time speed optimisation over a fixed path," *Int. Journal of Control*, vol. 87, no. 6, pp. 1297–1311, 2014.

[30] Y. Chen, M. Cutler, and J. P. How, "Decoupled multiagent path planning via incremental sequential convex programming," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5954–5961, IEEE, 2015.

[31] B. Zhou, F. Gao, J. Pan, and S. Shen, "Robust real-time uav replanning using guided gradient-based optimization and topological paths," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1208–1214, IEEE, 2020.

[32] Z. Zhang, Y. Zhong, J. Guo, Q. Wang, C. Xu, and F. Gao, "Auto filmer: Autonomous aerial videography under human interaction," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 784–791, 2022.

[33] L. Quan, L. Yin, T. Zhang, M. Wang, R. Wang, S. Zhong, X. Zhou, Y. Cao, C. Xu, and F. Gao, "Robust and efficient trajectory planning for formation flight in dense environments," *IEEE Transactions on Robotics*, vol. 39, no. 6, pp. 4785–4804, 2023.

[34] A. Walther and A. Griewank, "Getting started with adol-c.," *Combinatorial scientific computing*, vol. 1, no. 06.02, 2009.

[35] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, pp. 25–57, 2006.

[36] K. Konstantinidis, M. Alirezaei, and S. Grammatico, "Development of a detection and tracking of moving vehicles system for 2d lidar sensors," *Delft Universitu of Technology*, 2020.