Collaborative Quadruped Transportation in 3D Terrain with Constrained Diffusion

Williard Joshua Jose, Li Chen, Hao Zhang

***Abstract*— Recently, multi-robot systems have gained significant attention for their promise of scalable efficiency, reliability, and cost savings. A crucial capability is collaborative transportation, where a team of robots works together to transport a payload. However, key challenges remain, such as potential conflicts between team-level decisions and individual-level robot controls, team kinematic constraints imposed by the robot-payload coupling, and diverse obstacles encountered in 3D terrain. We present Collaborative Quadruped Transportation with Constrained Diffusion (CQTD), enabling a team of closely coupled quadruped robots to collaboratively transport a payload across 3D terrain. A diffusion-based upper level learns terrain-aware team-level trajectories satisfying team kinematic constraints due to the payload coupling, while a lower level optimizes velocity controls of individual robots satisfying collision and anisotropic velocity constraints. Experiments in high-fidelity simulations and on real-world quadruped robot teams demonstrate that CQTD outperforms baseline methods in challenging 3D terrain scenarios requiring closely-coupled collaboration between the quadruped robots.**

I. INTRODUCTION

Multi-robot systems are becoming an increasingly prominent area of robotics, with diverse real-world applications such as search and rescue [1], construction [2], manufacturing [3], and space exploration [4]. Using multiple smaller, cheaper robots instead of a single large, expensive robot results in improved operational efficiency, greater redundancy and reliability, and reduced deployment costs [5], [6]. Collaboration is a cornerstone of multi-robot systems, which is defined as a group of robots working closely and interdependently to achieve tasks that would be impossible for an individual robot to accomplish alone. Specifically, collaborative transportation is an essential capability within multi-robot collaboration, with the objective of enabling robots to work as a team to transport payloads to a designated goal position (Fig. 1).

Recent advancements in learning-based methods have significantly enhanced multi-robot collaboration capabilities. For example, deep neural networks (DNNs) have been implemented for multi-robot collision avoidance [7], [8], graph neural networks (GNNs) have demonstrated promise in decentralized multi-robot path planning [9]–[11], and deep reinforcement learning (DRL) has also been developed to perform collaborative multi-robot transportation, which generates motion plans for teams of quadrotors [12], [13] and quadrupeds [14]–[16]. However, several key challenges remain inadequately addressed. When multiple robots work together, conflicts can emerge between team-level decisions and individual robot controls. Previous methods have primarily focused on team-level motion planning and treat individual robot control as a separate task [17], [18]. Collaborative transportation can impose kinematic constraints when robots are coupled to a payload. Existing methods typically account only for the footprint of the payload when addressing collisions and represent the robots as point masses [8], [12], [19], which generally cannot address space-constrained scenarios.

The authors are with the Human-Centered Robotics Laboratory, University of Massachusetts Amherst, Amherst, MA 01002, USA. Email: {wjose,lchen0,hao.zhang}@umass.edu.

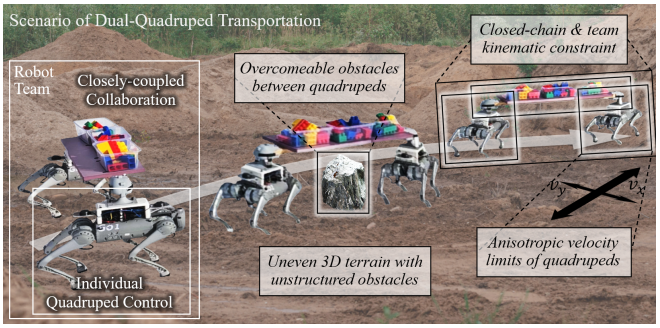

Fig. 1. A motivating scenario involves two quadruped robots collaborating closely to transport a payload across 3D terrain while avoiding obstacles that can be overcome between them.

To address these challenges, we introduce ***C**ollaborative **Q**uadruped **T**ransportation with Constrained **D**iffusion* (**CQTD**) that enables a team of two quadruped robots to collaboratively transport a payload to a goal position. CQTD simultaneously learns team-level collaboration decisions and individual-level robot controls under the unified mathematical framework of constrained bilevel optimization. At the upper level, a diffusion model learns to denoise robot team navigation trajectories to collaboratively move the payload to a designated goal position across 3D terrain with obstacles. At the lower level, CQTD optimizes the velocity controls of individual robots, enabling each robot to closely execute the team-level policy for collaboration, while satisfying kinematic and environment collision constraints and the robot's anisotropic velocity constraints.

Our contributions are listed below:

- This work enables a new multi-robot capability for a team of quadruped robots to collaboratively transport a payload across unstructured 3D outdoor terrain, while allowing the payload to overcome lower obstacles and pass through narrow alleys.
- We introduce the CQTD approach which simultaneously learns a team collaboration policy at the upper level and optimizes individual robot controls at the lower level,

while satisfying a set of constraints imposed by the team, individual robots, and the environment.

II. RELATED WORK

Multi-Robot Collaborative Transportation. A team of two or more robots can work together to collaboratively transport payloads that are either too large or too heavy for a single robot [20]–[22]. Previous approaches have studied collaborative transportation employing teams of unmanned ground vehicles (UGVs) [17], [18]. Traditional techniques used for multi-robot collaborative transportation are often based on model predictive control (MPC) or optimization [23]–[25], designed in order to enhance locomotion stability by treating the team as a single larger unit. However, these methods generally assume predefined trajectories for the team to transport the payload, which are generated by separate motion-planning techniques. Other tracks of research have implemented motion planning methods to plan navigational trajectories for the team to perform collaborative transportation, for example, using search-based planners [19], sampling-based planners [26], [27], and potential field planners [18]. More recent advancements have employed reinforcement learning-based planners that directly generate navigational trajectories for individual robots [12]. However, these previous approaches focus solely on collisions caused by the payload, neglecting individual robot collisions and their differing ground clearances. Additionally, when applied to quadruped robots, they fail to account for the anisotropic velocity limits specific to each robot in the team.

Diffusion Models for Robot Policy Learning. Diffusion models are a category of latent variable generative models that iteratively sample data from an underlying distribution through a denoising process. Recently, diffusion-based methods have attracted significant attention in robotics, where they have been applied to policy learning across a wide range of tasks in robotics, ranging from representing robot policies that denoise manipulator trajectories online [28], optimizing data-driven trajectories for motion planning [29], [30], and learning diverse robot behaviors [31], [32], but are limited to single robots or treat them as point masses during planning. Several approaches have explored diffusion-based policy learning for multi-agent systems such as human-robot collaboration [33], [34] and multi-robot motion planning [35], [36], but they do not adequately address the kinematic constraints imposed on the entire team and primarily focus on relatively controlled environments, failing to account for unstructured 3D environments.

III. APPROACH

We propose a novel approach, ***C**ollaborative **Q**uadruped **T**ransportation with Constrained **D**iffusion* (**CQTD**), to enable the multi-robot capability of closely-coupled collaborative transportation, which allows a team of quadruped robots to navigate complex 3D outdoor environments together and collaboratively transport a payload to a designated goal position with minimal time and distance traveled. CQTD develops a unified mathematical framework of constrained bilevel optimization to jointly learn team-level collaboration strategies and individual robot control policies. The upper level of CQTD uses a diffusion probabilistic model to generate navigational trajectories that guide the two quadruped robots in collaboratively transporting the payload. As the constraints of the upper level, CQTD's lower level optimizes the velocity controls of each quadruped robot, ensuring that the robots precisely follow the team-level collaboration policy. Constraints arising from the team kinematics, individual robot configurations, and interactions with the environment are also integrated into CQTD's mathematical formulation. We illustrate the CQTD approach in Figure 2.

Problem Definition. We represent the i-th robot using a floating 3D bounding box, parameterized by $\mathbf{b}_{r_i} = (\mathbf{x}_{r_i}, \mathbf{R}_{r_i}, \mathbf{d}_{r_i})$, where $\mathbf{x}_{r_i} = (x_{r_i}, y_{r_i}, z_{r_i}) \in \mathbb{R}^3$ is the robot's position in 3D space, $\mathbf{R}_{r_i} \in SO(3)$ represents its orientation as a 3D rotation matrix, and $\mathbf{d}_{r_i} \in \mathbb{R}^3$ represents the size (i.e., length, width, and height) of the bounding box. The quadruped robots are independently controlled with velocity commands in a decentralized fashion, represented for the i-th robot as $\mathbf{u}_{r_i} = (\mathbf{v}_{r_i}, \omega_{r_i})$, where $\mathbf{v}_{r_i} = (v_{x,r_i}, v_{y,r_i})$ and ω_{r_i} denote the linear and angular velocities of the robot, respectively.

Similarly, we model the payload as a floating 3D rectangular box, denoted by $\mathbf{b}_p = (\mathbf{x}_p, \mathbf{R}_p, \mathbf{d}_p)$, where $\mathbf{x}_p = (x_p, y_p, z_p) \in \mathbb{R}^3$ denotes its 3D position, $\mathbf{R}_p \in SO(3)$ defines its 3D orientation, and $\mathbf{d}_p \in \mathbb{R}^3$ represents its dimensions. Furthermore, we use $\mathbf{x}_{start} \in \mathbb{R}^3$ and $\mathbf{x}_{goal} \in \mathbb{R}^3$ to represent the start and goal positions of the payload in 3D space, respectively. The two ends of the payload are mounted on the backs of the two quadruped robots using ball joints, which constrain the relative position between the robots and the payload while allowing rotations. We model this *kinematic constraint* in 3D space using a function $f : \mathbf{b}_p \mapsto (\mathbf{x}_{r_1}, \mathbf{x}_{r_2})$, and also imposes a *closed-chain constraint* for the team, which requires both quadrupeds to remain in contact with the ground surface. The constraint can be represented as $\mathbf{T}_p^{r_1}\mathbf{x}_{r_1} = \mathbf{T}_p^{r_2}\mathbf{x}_{r_2} = \mathbf{x}_p$, where $\mathbf{T}_p^{r_1} \in SE(3)$ and $\mathbf{T}_p^{r_2} \in SE(3)$ are the transformation matrices from robots #1 and #2 to the payload's center of mass.

Outdoor field environments typically feature uneven terrain and irregular obstacles preventing the robot team from passing through, which we define as the *collision constraints*. To ensure safety while navigating, the quadrupeds must satisfy the collision constraints to avoid collisions with the terrain and obstacles. Mathematically, we represent the terrain using an elevation map $\mathbf{H} : (x, y) \mapsto z$, where z is the height of the uneven terrain at position (x, y). Obstacles are represented in the elevation map as an increase in terrain height at their respective locations. We define the i-th robot as free from collision with the environment when its bounding box remains above the terrain and obstacles, i.e., $b_z > \mathbf{H}(b_x, b_y), \forall\ (b_x, b_y, b_z) \in \mathbf{b}_{r_i}$, where (b_x, b_y, b_z) is any point from the robot's bounding box $\mathbf{b}_{r_i}$. We similarly define collision constraints with respect to $\mathbf{b}_p$ for the payload.

Our objective is to determine a sequence of robot velocity

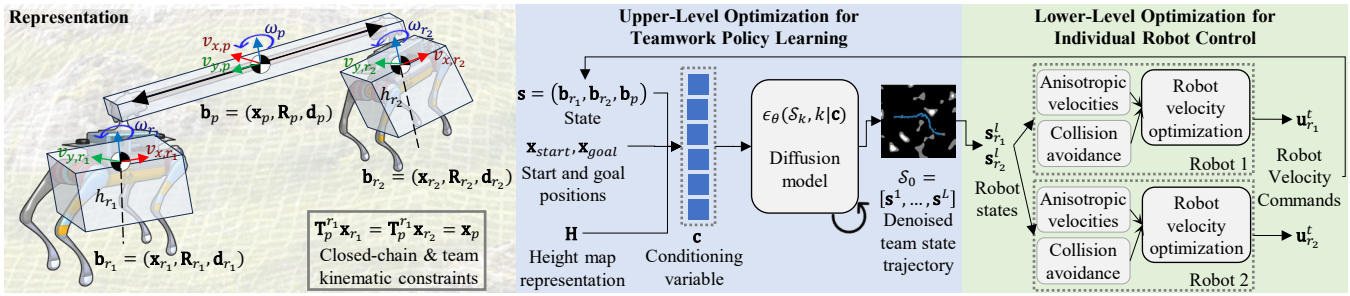

Fig. 2. Overview of CQTD. Collaborative transportation is formulated as constrained bilevel optimization: the upper level employs diffusion to learn a team-level trajectory to transport the payload, while the lower level computes individual robot velocity controls satisfying closed-chain kinematic constraints, collision avoidance constraints, and anisotropic robot velocity constraints.

commands that minimize the time and distance required for the robots to transport the payload from a start position $\mathbf{x}_{start}$ to a goal position $\mathbf{x}_{goal}$ while respecting kinematic and collision constraints.

Upper-Level Optimization for Teamwork Policy Learning. We introduce a diffusion probabilistic model at the upper level of CQTD to learn the collaboration between the quadrupeds, which enables them to transport the payload to the goal position. Formally, we define the team state as a set of bounding boxes that denote both the robots and payload, i.e., $\mathbf{s} = (\mathbf{b}_{r_1}, \mathbf{b}_{r_2}, \mathbf{b}_p)$, where each $\mathbf{b}$ consists of the position $\mathbf{x}$, orientation $\mathbf{R}$, and dimensionality $\mathbf{d}$ of the bounding box in 3D space. We further represent a trajectory of team states as $\mathcal{S} = [\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^L]$ that includes L team states on the trajectory. Then, we model the team trajectory probabilistically as a conditional distribution $p(\mathcal{S}|\mathbf{c})$, where $\mathbf{c} = (\mathbf{H}, \mathbf{x}_{start}, \mathbf{x}_{goal})$ encodes the height map representation $\mathbf{H}$, start position $\mathbf{x}_{start}$, and goal position $\mathbf{x}_{goal}$. The conditional distribution $p(\mathcal{S}|\mathbf{c})$ resides in a high-dimensional space, which makes direct modeling and sampling intractable. To explicitly address this challenge, we propose to leverage the expressive power of diffusion models to learn this distribution as a parameterized model.

The diffusion model is based upon the probabilistic diffusion process [37], with Gaussian noise progressively added to the team state trajectory $\mathcal{S}_0$ during the forward diffusion process to obtain noisy trajectories $\mathcal{S}_1, \ldots, \mathcal{S}_{K-1}, \mathcal{S}_K$ using a Gaussian function q, while maintaining features of the original trajectory $\mathcal{S}_0$. This forward diffusion process can be expressed by $q(\mathcal{S}_k|\mathcal{S}_{k-1}, k) = \mathcal{N}(\mathcal{S}_k, \sqrt{1-\beta_k}\mathcal{S}_{k-1}, \beta_k \mathbf{I})$, where β_k is determined by the diffusion noise variance schedule, and $k \in \{1, \ldots, K\}$. The reverse diffusion model then reverses the forward process via predicting the Gaussian noise at each step k utilizing a deep neural network $\epsilon_\theta(\mathcal{S}_k, k)$ parameterized by θ. Starting with a standard Gaussian prior $p(\mathcal{S}_K) = \mathcal{N}(0, I)$, the denoising process iteratively transforms the noisy trajectory $\mathcal{S}_K$ into a noise-free trajectory $\mathcal{S}_0$. Since the navigational trajectories of the team are influenced by the environment, as well as the team's start and goal positions, we design the denoised trajectories to be conditioned on the height map, start position, and goal position. The conditional denoising process is defined as $\mathcal{S}_{k-1} = \alpha_k\left(\mathcal{S}_k - \gamma_k \epsilon_\theta(\mathcal{S}_k, k|\mathbf{c})\right) + \sigma_k$, where α_k, γ_k, and σ_k are parameters in this reverse process. To learn this diffusion model, we minimize the mean squared error between the predicted noise $\epsilon_\theta(\mathcal{S}_k, k|\mathbf{c})$ and the actual noise added up to the k-th step, expressed by $\mathcal{L} = \mathbb{E}_{\mathbf{c},\mathcal{S}_0}\|\epsilon_k - \epsilon_\theta(\mathcal{S}_0 + \epsilon_k, k|\mathbf{c})\|_2^2$, where ϵ_k is a randomly sampled noise with variance σ_k.

Then, the upper-level optimization of CQTD can be mathematically expressed by:

$$\min_\theta \mathbb{E}_{\mathbf{c},\mathcal{S}_0}\|\epsilon_k - \epsilon_\theta(\mathcal{S}_0 + \epsilon_k, k|\mathbf{c})\|_2^2 \tag{1}$$

$$\text{s.t. } (\mathbf{x}_{r_1}^t, \mathbf{x}_{r_2}^t) = f(\mathbf{b}_p^t) \tag{2}$$

$$\mathbf{T}_p^{r_1}\mathbf{x}_{r_1}^t = \mathbf{T}_p^{r_2}\mathbf{x}_{r_2}^t = \mathbf{x}_p^t \tag{3}$$

$$b_z^t > \mathbf{H}(b_x^t, b_y^t) \;\forall\; (b_x^t, b_y^t, b_z^t) \in \mathbf{b}_j^t, j \in \{p, r_1, r_2\} \tag{4}$$

The objective function in Eq. (1) aims to learn the parameters θ of the diffusion model used to sample the maximum likelihood of $p(\mathcal{S}|\mathbf{c})$ in order to generate navigational trajectories for the team of robots to collaboratively transport the payload to the goal position, while satisfying the team kinematic and closed-chain constraints in Eqs. (2) and (3), respectively, and the environment constraints for the team and payload to avoid collisions in Eq. (4).

Lower-Level Optimization for Individual Robot Control. The goal of our CQTD's lower-level optimization is to generate a sequence of L velocity controls $\mathcal{U}_{r_i} = [\mathbf{u}_{r_i}^1, \mathbf{u}_{r_i}^2, \ldots \mathbf{u}_{r_i}^L]$ for each individual robot, enabling them to closely follow the upper level team state trajectory and reach their goal positions as fast as possible, while satisfying collision and velocity constraints.

Given the l-th team state $\mathbf{s}^l \in \mathcal{S}$ from the trajectory of team states $\mathcal{S} = [\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^L]$ planned by our CQTD's upper-level optimization, we extract and define the bounding box poses of the i-th robot as its state $\mathbf{s}_{r_i}^l = (\mathbf{x}_{r_i}^l, \mathbf{R}_{r_i}^l) \in \mathbf{s}^l$. When the robot takes an action denoted by a velocity command $\mathbf{u}_{r_i}^l$ for a time period $\tau_{r_i}^l$, this robot transitions from its current state $\mathbf{s}_{r_i}^l$ to the next robot state $\mathbf{s}_{r_i}^{l+1}$ through $\mathbf{s}_{r_i}^{l+1} = \mathbf{s}_{r_i}^l + \int \mathbf{u}_{r_i}^t \mathrm{d}t$. The objective function of the lower level, denoted by a squared ℓ_2-loss, is designed to ensure that the actual states $(\mathbf{s}_{r_i}^l)^{act}$ of the i-th quadruped robot closely follow its planned states $(\mathbf{s}_{r_i}^l)^{plan} = \mathbf{s}_{r_i}^l$ generated at the upper level. Furthermore, the lower level of CQTD aims to minimize the total travel time to encourage each robot to move faster, which can be accomplished through minimizing $\sum_{l=1}^L \tau_{r_i}^l$. Non-holonomic ground robots, including quadrupedal robots, are subject to *anisotropic velocity constraints*, where the robot's maximum allowable velocities

differ for movements along the x-axis and y-axis. To explicitly model such constraints in CQTD's lower level, we define the maximum anisotropic velocities of each quadrupedal robot as $\mathbf{v}_{max} = (v_{x,max}, v_{y,max})$, where $v_{x,max} > v_{y,max} > 0$, and require linear velocities to satisfy these anisotropic velocity constraints $\left(|v^t_{x,r}|, |v^t_{y,r}|\right) \leq \mathbf{v}_{max}$. Incorporating the minimization objective and constraints, we mathematically formulate the lower level of CQTD as:

$$\min_{\mathcal{U}_{r_i}} \sum_{l=1}^{L} \left(\tau^l_{r_i} + \lambda \|(\mathbf{s}^l_{r_i})^{act} - (\mathbf{s}^l_{r_i})^{plan}\|_2^2\right), \forall i \in \{1,2\} \quad (5)$$

$$\text{s.t.} \quad \left(|v^t_{x,r_i}|, |v^t_{y,r_i}|\right) \leq \mathbf{v}_{max}, \forall i \in \{1,2\} \quad (6)$$

$$b^t_z > \mathbf{H}(b^t_x, b^t_y) \ \forall \ (b^t_x, b^t_y, b^t_z) \in \mathbf{b}^t_j, j \in \{r_1, r_2\} \quad (7)$$

where λ is a hyperparameter to balance the two loss functions. Besides the anisotropic velocity constraints, each individual robot must satisfy the collision constraints for obstacle avoidance as modeled in Eq. (7).

Constrained Bilevel Optimization for Diffusion Learning. We integrate the upper-level and lower-level objectives into the unified mathematical framework of constrained bilevel optimization, which results in our CQTD approach, defined as:

$$\min_{\theta} \mathbb{E}_{\mathbf{c},\mathcal{S}_0} \|\epsilon_k - \epsilon_\theta(\mathcal{S}_0 + \epsilon_k, k|\mathbf{c})\|_2^2 \quad (8)$$

$$\text{s.t.} \ \underset{\mathcal{U}_{r_i}}{\text{argmin}} \sum_{l=1}^{L} \left(\tau^l_{r_i} + \lambda \|(\mathbf{s}^l_{r_i})^{act} - (\mathbf{s}^l_{r_i})^{plan}\|_2^2\right), \forall i \in \{1,2\} \quad (9)$$

$$(\mathbf{x}^t_{r_1}, \mathbf{x}^t_{r_2}) = f(\mathbf{b}^t_p) \quad (10)$$

$$\mathbf{T}^{r_1}_p \mathbf{x}^t_{r_1} = \mathbf{T}^{r_2}_p \mathbf{x}^t_{r_2} = \mathbf{x}^t_p \quad (11)$$

$$\left(|v^t_{x,r_i}|, |v^t_{y,r_i}|\right) \leq \mathbf{v}_{max}, \forall i \in \{1,2\} \quad (12)$$

$$b^t_z > \mathbf{H}(b^t_x, b^t_y) \ \forall \ (b^t_x, b^t_y, b^t_z) \in \mathbf{b}^t_j, j \in \{p, r_1, r_2\} \quad (13)$$

The upper level in Eq. (8) learns a trajectory of team states $\mathcal{S}$ as the collaboration policy for the quadrupeds to transport the payload together, while CQTD's lower level in Eq. (9), as a constraint of the upper-level optimization, optimizes velocity controls for individual robots to precisely execute the team policy in a decentralized manner. Within the optimization formulation, CQTD simultaneously integrates four constraints. The constraint at the upper level on team kinematics is modeled by Eq. (10), which arises from physical connections between the payload and the robots. The closed-chain constraint at the upper level, modeled by Eq. (11), ensures that the robots remain on the ground surface. The constraints on anisotropic velocities at the lower level in Eq. (12) reflect the different capabilities of the quadruped robots when moving in different directions. The collision constraints in Eq. (13) must be satisfied both by each individual robot at the lower level and the payload carried by the team at the upper level of CQTD.

In order to solve the formulated constrained bilevel optimization problem, we simultaneously train the diffusion probabilistic model to generate trajectories of the robot team $\mathcal{S}$ at the upper level in Eq. (8) and optimize individual robot velocities $\mathcal{U}_r$ at the lower level in Eq. (9). Given a dataset of collaborative transportation $\mathcal{D} = \{\mathbf{c}, \mathcal{S}^*\}_{i=1}^N$, which contains N samples that are collected using our new multi-robot simulation (presented in Sec. IV) across varying height maps $\mathbf{H}$ of different terrain, as well as the robot team's start positions $\mathbf{x}_{start}$, goal positions $\mathbf{x}_{goal}$, and trajectories $\mathcal{S}^*$, we perform imitation learning to generate a trajectory of team states. The upper level Eq. (8) optimizes the diffusion model parameters θ to minimize the diffusion loss with respect to $\mathcal{S}^*$ and $\mathbf{c}$ in the dataset and is solved by training the diffusion model using gradient descent based upon the Adam optimizer [38], [39]. The lower-level objective is optimized during the training phase by minimizing the error between the learned individual robot velocities and the velocities from the imitation samples within the dataset. During execution, this is implemented using a PID-based local planner that tracks the robot team states from $\mathcal{S}$, and individual robot velocities are scaled down to within velocity limits to prevent robot and payload collisions.

TABLE I

QUANTITATIVE RESULTS OF CQTD IN SIMULATION.

Method	Success Rate	Planning Time (s)	Path Length (m)
A* [40]	76.40%	4.31	14.92
RRT* [41]	67.40%	0.21	18.66
DQN [12]	13.40%	0.06	15.86
CQTD (ours)	86.20%	0.76	17.00

IV. EXPERIMENTS

Experimental Setups. We utilize two Unitree Go1 quadruped robots using the Champ locomotion controller [42], [43] running on Robot Operating System (ROS) 1. We create a synthetically generated dataset to train our CQTD approach across 100,000 different height maps with 10 varying start and goal positions each. Specifically, the height maps $\mathbf{H}$ are generated using Perlin noise [44] for realistic terrain and are encoded into a height map representation using a ResNet-18 network [45]. The upper-level optimal robot team trajectories $\mathcal{S}^*$ are solved by a search-based motion planner considering lower-level collision constraints and anisotropic velocity limits of the quadrupeds.

We compare CQTD against three baseline approaches: (1) **A*** [40], a graph-based pathfinding algorithm based on a weighted graph of the configuration space, (2) **RRT*** [41], a sampling-based method that explores the robot's configuration space by constructing a space-filling tree, and (3) **DQN** [12], a learning-based method using deep Q-network (DQN) for cooperative object transportation.

Our CQTD is quantitatively evaluated and compared against the baseline methods under three metrics: (1) **success rate**, the ratio of completed collaborative tasks to total evaluations, (2) **planning time**, which is the total time cost for robot team's actions inference, and (3) **total path length**, which is the total distance traveled by the payload during collaborative transportation. Both planning time and total path length are measured only for successful cases.

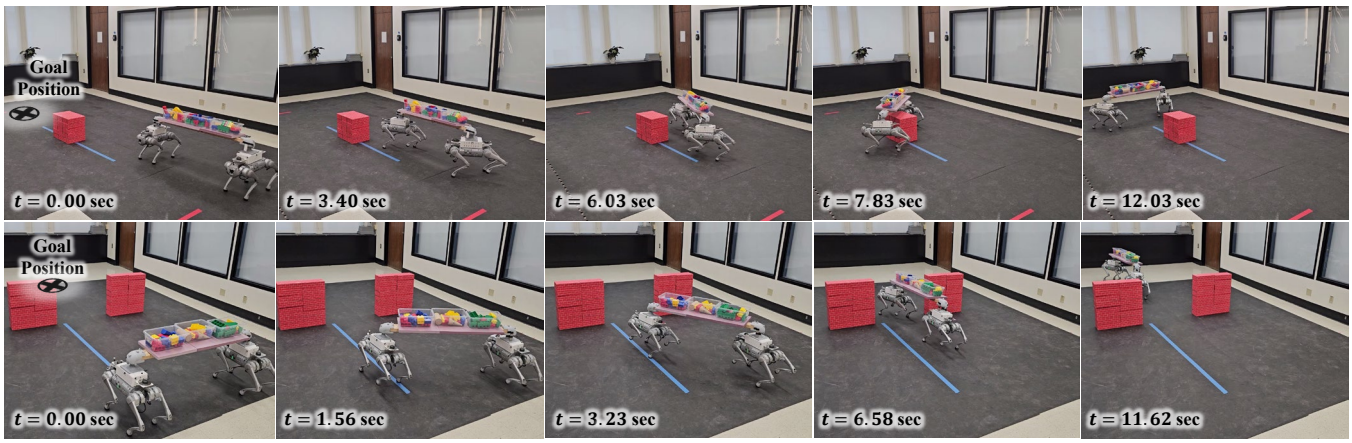

Fig. 3. Two quadruped robots collaboratively transport brick payloads as a closely-coupled team in an indoor environment (a) over a block obstacle and (b) through a narrow alley between block walls.

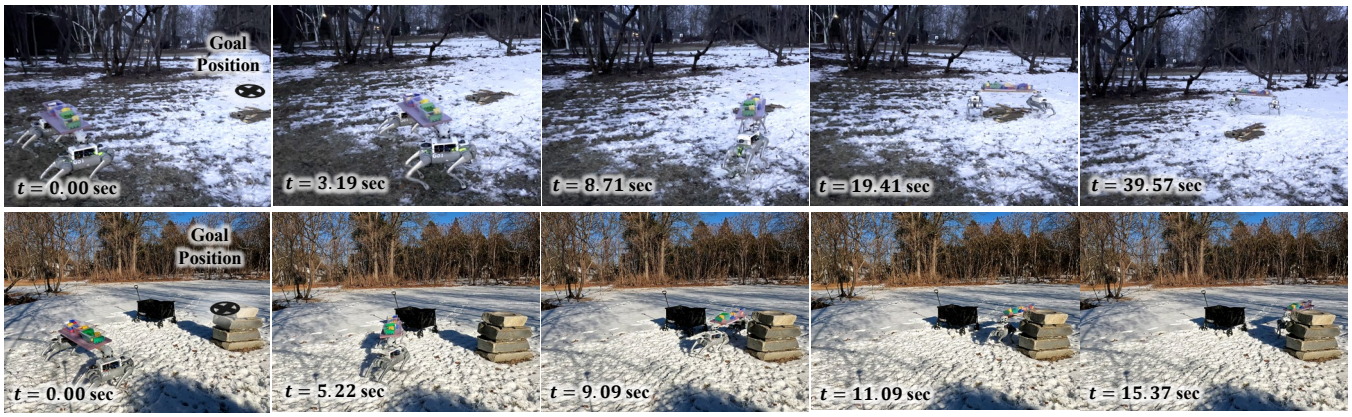

Fig. 4. A closely-coupled team of two quadruped robots collaboratively transports brick payloads across an outdoor environment with unstructured 3D terrain (a) over tree branches, and (b) between a pile of rocks and a cart.

Quantitative and Qualitative Results on Simulation. We evaluate our CQTD approach on the test split of the synthetically generated terrain dataset. We collect metrics on 500 unique terrain environments with varying start and goal positions. Our quantitative results can be seen in Table I. CQTD has the highest success rate among the baseline methods, outperforming them by at least ten percentage points. Although RRT* has the fastest planning time (due to its probabilistically complete formulation), the path length it produces is the longest and its success rate is not very high. A* has the shortest path length because its solutions (should it find some) are guaranteed shortest; however, its search-based nature significantly slows down its planning time due to the exponential time complexity. Also, A* can also suffer from discretization errors when optimizing search speed. DQN, the only comparable learning-based baseline available in recent work, performs poorly due to the difficulty in learning a generalized policy across varying environments.

Qualitative Results on Real-World Robots. Additionally, two Unitree Go1 quadruped robots are deployed in both indoor and outdoor real-world scenarios, where CQTD fully autonomously controls the robots. In the indoor environment, CQTD successfully finishes the *overcome obstacle* scenario, where the payload passes over the obstacle while the robots walk side-by-side, illustrated in Fig. 3(a). CQTD also executes the *pass narrow alley* scenario in Fig. 3(b), where the robots switch from side-by-side to instead form a line to enter the narrow area. In the outdoor environment, we illustrate the same *overcome obstacle* and *pass narrow alley* scenarios in Fig. 4(a) and Fig. 4(b).

V. CONCLUSION

In this work, we introduced CQTD as a novel constrained bilevel optimization approach designed to enable collaborative payload transportation across complex 3D terrain. CQTD uses a diffusion-based learning method to generate robot team trajectories at the upper level, while simultaneously generating individual robot velocity commands at the lower level. At the upper level, it considers closed-chain and team kinematic constraints, as well as collision avoidance constraints for both the robots and the payload, while CQTD's lower level accounts for collision avoidance constraints and anisotropic velocity limitations for each individual quadruped. In order to validate its performance, we conducted comprehensive experiments to evaluate CQTD through Gazebo simulations and physical quadruped robot teams collaboratively transporting a payload in indoor and outdoor environments. Experimental results demonstrate that our CQTD approach enables new multi-robot capabilities for collaborative transportation through

robot learning, and it outperforms baseline methods across multiple metrics.

REFERENCES

[1] J. P. Queralta, J. Taipalmaa, B. Can Pullinen, V. K. Sarker, T. Nguyen Gia, H. Tenhunen, M. Gabbouj, J. Raitoharju, and T. Westerlund, "Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision," *IEEE Access*, vol. 8, pp. 191 617–191 643, 2020.

[2] M. Krizmancic, B. Arbanas, T. Petrovic, F. Petric, and S. Bogdan, "Cooperative aerial-ground multi-robot system for automated construction tasks," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 798–805, 2020.

[3] Z. Nie and K.-C. Chen, "Predictive path coordination of collaborative transportation multirobot system in a smart factory," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–14, 2024.

[4] H. Farivarnejad, A. S. Lafmejani, and S. Berman, "Fully decentralized controller for multi-robot collective transport in space applications," in *IEEE Aerospace Conference*, 2021.

[5] L. V. Nguyen, "Swarm intelligence-based multi-robotics: A comprehensive review," *AppliedMath*, vol. 4, no. 4, pp. 1192–1210, 2024.

[6] C. Ju, J. Kim, J. Seol, and H. I. Son, "A review on multirobot systems in agriculture," *Computers and Electronics in Agriculture*, vol. 202, p. 107336, 2022.

[7] G. Shi, W. Hönig, Y. Yue, and S.-J. Chung, "Neural-swarm: Decentralized close-proximity multirotor control using learned interactions," in *IEEE International Conference on Robotics and Automation*, 2020.

[8] B. Rivière, W. Hönig, Y. Yue, and S.-J. Chung, "Glas: Global-to-local safe autonomy synthesis for multi-robot motion planning with end-to-end learning," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4249–4256, 2020.

[9] Q. Li, F. Gama, A. Ribeiro, and A. Prorok, "Graph neural networks for decentralized multi-robot path planning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020.

[10] Y. Jo and H. I. Son, "A gnn-based decentralized path planning for agricultural robot team: Work in progress," in *IEEE International Conference on Control, Automation and Systems*, 2024.

[11] V. D. Sharma, L. Zhou, and P. Tokekar, "D2coplan: A differentiable decentralized planner for multi-robot coverage," in *IEEE International Conference on Robotics and Automation*, 2023.

[12] L. Zhang, Y. Sun, A. Barth, and O. Ma, "Decentralized control of multi-robot system in cooperative object transportation using deep reinforcement learning," *IEEE Access*, vol. 8, pp. 184 109–184 119, 2020.

[13] H. Zhu, F. M. Claramunt, B. Brito, and J. Alonso-Mora, "Learning interaction-aware trajectory predictions for decentralized multi-robot motion planning in dynamic environments," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2256–2263, 2021.

[14] H. Zhu, S. Yang, W. Wang, X. He, and N. Ding, "Cooperative transportation of tether-suspended payload via quadruped robots based on deep reinforcement learning," in *IEEE International Conference on Robotics and Biomimetics*, 2023.

[15] Y. Ji, B. Zhang, and K. Sreenath, "Reinforcement learning for collaborative quadrupedal manipulation of a payload over challenging terrain," in *IEEE International Conference on Automation Science and Engineering*, 2021.

[16] Y. Feng, C. Hong, Y. Niu, S. Liu, Y. Yang, W. Yu, T. Zhang, J. Tan, and D. Zhao, "Learning multi-agent loco-manipulation for long-horizon quadrupedal pushing," *arXiv preprint arXiv:2411.07104*, 2024.

[17] Q. Liu, Z. Nie, Z. Gong, and X.-J. Liu, "An omnidirectional transportation system with high terrain adaptability and flexible configurations using multiple nonholonomic mobile robots," *IEEE Robotics and Automation Letters*, vol. 8, no. 9, pp. 6060–6067, 2023.

[18] H. Cohen, S. Hacohen, N. Shvalb, and O. Medina, "Decentralized motion planning for load carrying and manipulating by a robotic pack," *IEEE Access*, vol. 11, pp. 16 557–16 566, 2023.

[19] C. Yang, G. N. Sue, Z. Li, L. Yang, H. Shen, Y. Chi, A. Rai, J. Zeng, and K. Sreenath, "Collaborative navigation and manipulation of a cable-towed load by multiple quadrupedal robots," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10 041–10 048, 2022.

[20] Z. Yan, N. Jouandeau, and A. A. Cherif, "A survey and analysis of multi-robot coordination," *International Journal of Advanced Robotic Systems*, vol. 10, no. 12, p. 399, 2013.

[21] 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.

[22] H. Farivarnejad and S. Berman, "Multirobot control strategies for collective transport," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, no. 1, pp. 205–219, 2022.

[23] A. Tagliabue, M. Kamel, R. Siegwart, and J. Nieto, "Robust collaborative object transportation using multiple mavs," *The International Journal of Robotics Research*, vol. 38, no. 9, pp. 1020–1044, 2019.

[24] J. Wehbeh, S. Rahman, and I. Sharf, "Distributed model predictive control for uavs collaborative payload transport," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020.

[25] Z. Wang, S. Singh, M. Pavone, and M. Schwager, "Cooperative object transport in 3d with multiple quadrotors using no peer communication," in *IEEE International Conference on Robotics and Automation*, 2018.

[26] P. Yu and D. V. Dimarogonas, "Distributed motion coordination for multirobot systems under ltl specifications," *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 1047–1062, 2022.

[27] H. Zhang, H. Song, W. Liu, X. Sheng, Z. Xiong, and X. Zhu, "Hierarchical motion planning framework for cooperative transportation of multiple mobile manipulators," *arXiv preprint arXiv:2208.08054*, 2022.

[28] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," *The International Journal of Robotics Research*, p. 02783649241273668, 2023.

[29] J. Carvalho, A. T. Le, M. Baierl, D. Koert, and J. Peters, "Motion planning diffusion: Learning and planning of robot motions with diffusion models," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2023.

[30] J. Liu, M. Stamatopoulou, and D. Kanoulas, "Dipper: Diffusion-based 2d path planner applied on legged robots," in *IEEE International Conference on Robotics and Automation*, 2024.

[31] Z. Li, R. Krohn, T. Chen, A. Ajay, P. Agrawal, and G. Chalvatzaki, "Learning multimodal behaviors from scratch with diffusion policy gradient," *arXiv preprint arXiv:2406.00681*, 2024.

[32] G. Yan, Y.-H. Wu, and X. Wang, "Dnact: Diffusion guided multi-task 3d policy learning," *arXiv preprint arXiv:2403.04115*, 2024.

[33] T. Yoneda, L. Sun, G. Yang, B. C. Stadie, and M. R. Walter, "To the noise and back: Diffusion for shared autonomy," in *Robotics: Science and Systems*, 2023.

[34] E. Ng, Z. Liu, and M. Kennedy, "Diffusion co-policy for synergistic human-robot collaborative tasks," *IEEE Robotics and Automation Letters*, vol. 9, no. 1, pp. 215–222, 2024.

[35] Y. Shaoul, I. Mishani, S. Vats, J. Li, and M. Likhachev, "Multi-robot motion planning with diffusion models," *arXiv preprint arXiv:2410.03072*, 2024.

[36] Z. Wu, S. Ye, M. Natarajan, and M. C. Gombolay, "Diffusion-reinforcement learning hierarchical motion planning in adversarial multi-agent games," *arXiv preprint arXiv:2403.10794*, 2024.

[37] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Advances in Neural Information Processing Systems*, 2020.

[38] D. Kinga, J. B. Adam *et al.*, "A method for stochastic optimization," in *International conference on learning representations*, 2015.

[39] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in neural information processing systems*, 2019.

[40] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.

[41] S. Karaman and E. Frazzoli, "Incremental sampling-based algorithms for optimal motion planning," in *Robotics: Science and Systems*, 2010.

[42] "chvmp/champ," 2020.

[43] J. Lee, "Hierarchical controller for highly dynamic locomotion utilizing pattern modulation and implementation on the MIT Cheetah robot," Master's thesis, Massachusetts Institute of Technology, 2013.

[44] K. Perlin, "Improving noise," in *Conference on computer graphics and interactive techniques*, 2002.

[45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE/CVF Conference on computer vision and pattern recognition*, 2016.