

# Learning Decentralized Flocking Controllers with Spatio-Temporal Graph Neural Network

Siji Chen<sup>1</sup>, Yanshen Sun<sup>1</sup>, Peihan Li<sup>2</sup>, Lifeng Zhou<sup>2</sup>, Chang-Tien Lu<sup>1</sup>

**Abstract**—Recently a line of research has delved into the use of graph neural networks (GNNs) for decentralized control in swarm robotics. However, it has been observed that relying solely on the states of immediate neighbors is insufficient to imitate a centralized control policy. To address this limitation, prior studies proposed incorporating  $L$ -hop delayed states into the computation. While this approach shows promise, it can lead to a lack of consensus among distant flock members and the formation of small clusters, consequently failing cohesive flocking behaviors. Instead, our approach leverages spatiotemporal GNN, named STGNN that encompasses both spatial and temporal expansions. The spatial expansion collects delayed states from distant neighbors, while the temporal expansion incorporates previous states from immediate neighbors. The broader information gathered from both expansions results in more effective and accurate predictions. We develop an expert algorithm for controlling a swarm of robots and employ imitation learning to train our decentralized STGNN model based on the expert algorithm. We simulate the proposed STGNN approach in various settings, demonstrating its decentralized capacity to emulate the global expert algorithm. Further, we implemented our approach to achieve cohesive flocking, leader following, and obstacle avoidance by a group of Crazyflie drones. The performance of STGNN underscores its potential as an effective and reliable approach for achieving cohesive flocking, leader following, and obstacle avoidance tasks.

## I. INTRODUCTION

Flocking is a collective behavior observed in groups of animals or autonomous agents, such as birds, fish, or artificial robots, where individuals move together in a coordinated manner. In a flock, each robot follows simple rules based on local information to achieve a common group objective [1]. Multi-robot systems based on flocking models exhibit self-organization and goal-directed behaviors, making them suitable for various applications, including automated parallel delivery, sensor network design, and search and rescue operations [2]. Flocking is typically modeled as a consensus or alignment problem, aiming to ensure that all robots in the group eventually agree on their states [3]. Classical methods such as those proposed by Tanner [4] and Olfati-Saber [3], define rules and constraints governing the position, speed, and acceleration of the robots. However, these methods heavily rely on parameter tuning and are limited to predefined scenarios. In contrast, learning-based methods spontaneously explore complex patterns and adapt

their parameters through training, providing more flexibility and adaptability compared to classical approaches.

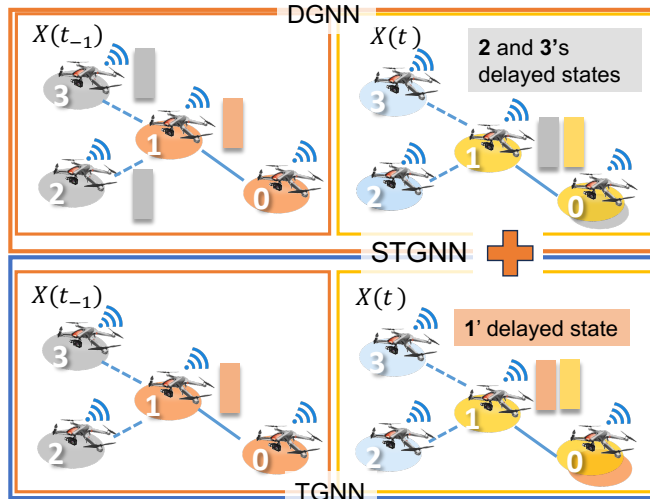


Fig. 1. Comparison of the state information used by DGNN (top) and TGNN (bottom): DGNN propagates  $t - 1$  states of robots 2 and 3 along with robot 1's current state at  $t$  to robot 0. TGNN propagates  $t - 1$  and  $t$  states of robot 1 to robot 0. STGNN merges both state information, resulting in superior predictive power compared to either method alone.

There are primarily two research directions in learning-based methods. One approach focuses on imitation learning, as demonstrated by Tolstaya et al. [5], Kortvelesy et al. [6], Zhou et al. [7], and Lee et al. [8]. The other approach involves multi-robot deep reinforcement learning (MADRL), as explored in the studies of Yan et al. [9] and Xiao et al. [10]. MADRL is particularly useful when labels are unavailable, but it presents its own set of challenges, including the demand for an extensive volume of training data and limitations in generalizing to new and unencountered scenarios [11]. In this work, we choose to utilize imitation learning due to the availability of an expert policy that has proven to be effective for our task [4], [5], [12].

Recent research in this direction adopts a graph-based approach to represent robots and their interactions leveraging Graph Neural Networks (GNNs) [13], [14] for modeling and analyzing flock dynamics. This approach shows promise in addressing flocking tasks by harnessing the power of graph-based representations and neural networks. Specifically, studies such as Tolstaya et al. [5], Kortvelesy et al. [6], Zhou et al. [7], and Lee et al. [8] utilize a technique called "delayed state" to incorporate the information from the  $l$ -step-before states of a robot's  $l$ -hop neighbors, where  $l = 1, 2, \dots, L$  [5].

<sup>2</sup>Siji Chen, Yanshen Sun, and Chang-Tien Lu are with the Department of Computer Science, Virginia Tech, Falls Church, VA 22043, USA. Email: {sijic, yansh93, ctlu}@vt.edu.

<sup>2</sup>Peihan Li and Lifeng Zhou are with the Department of Electrical and Computer Engineering, Drexel University, Philadelphia, PA 19104, USA. Email: {p1525, lz457}@drexel.edu.

Henceforth, we identify this type of model as the “delayed graph neural network” (DGNN). DGNN enables the learning of spatially extended representations in the local network. However, it overlooks the influence of a robot’s historical states and the historical states of its neighbors, thereby neglecting the temporal sequence of flock movement. Differently, the spatial and temporal expansion of STGNN enables it to gather information from both spatial and temporal dimensions. We illustrate the distinct information collected by DGNN and temporal GNN (TGNN), as well as their combined information by STGNN in Figure 1. With this insight, we make the following primary contributions:

- **Design an STGNN-based imitation learning framework for decentralized flocking control.** STGNN enables effective information fusion by integrating delayed states from distant neighbors and previous states from immediate neighbors. To the best of our knowledge, we are the first to design an STGNN-based learning framework for multi-robot flocking with leader following and obstacle avoidance.
- **Develop a centralized expert algorithm for flocking, with leader following and obstacle avoidance.** Finding an expert algorithm is the key to success in imitation learning. We develop an expert algorithm that provides effective control over a large swarm of robots based on Tanner’s [4] and Olfati-Saber’s work [3]. To the best of our knowledge, we are the first to offer a complete global expert algorithm, capable of flocking with both leader following and obstacle avoidance.
- **Conduct extensive evaluations of STGNN.** We comprehensively evaluate STGNN by comparing it to DGNN and TGNN and testing STGNN with varying history horizons. The results demonstrate STGNN’s effectiveness and superior performance in completing complex flocking tasks. Additionally, we implement STGNN into a group of Crazyflie drones to achieve flocking with obstacle avoidance.

## II. PROBLEM FORMULATION

Consider a collection of  $N$  robots that are identical and possess the same capabilities such as maximum acceleration  $U_{\max}$ , maximum velocity  $V_{\max}$ , and communication range  $R_c$ . Each robot  $i$  can be uniquely identified by its position  $\mathbf{p}_i$  and its velocity  $\mathbf{v}_i$ . The task for our learning model is to compute the control input  $\mathbf{u}_i$ , i.e., acceleration for each robot, based on the current state by itself and its neighbors, represented by the position and velocity  $(\mathbf{p}, \mathbf{v})$ .

In this paper, we aim to solve three problems as a whole: robot flocking, leader following, and flocking with obstacle avoidance. Each robot is required to avoid collisions with other robots, follow and maintain proximity to a virtual leader, and navigate around obstacles. Specifically, flocking is to maintain a consistent distance from other robots while synchronizing their movements. The leader following asks the robot swarm to track one or more leader(s) during flocking. We opt for a single virtual leader for the entire swarm, as the primary objective of flocking is to achieve

consensus, and introducing multiple leaders could potentially violate this objective. It is important to note that the virtual leader differs from other robots in that it does not need to avoid collisions. Instead, it represents a predefined trajectory known to all robots. Finally, the obstacle avoidance requests the robot swarm to achieve the two tasks above while avoiding crashing into obstacles.

## III. METHODOLOGY

In this research, we employ imitation learning to train STGNN. Specifically, we train an expert model whose outputs serve as labels for training our STGNN model. The expert assumes that the strategy provider possesses real-time information of all robots (i.e., centralized communication), which is impracticable in reality. In contrast, STGNN considered real-world scenarios, where each robot makes decisions by itself with local communication and delayed information (i.e., decentralized communication). Due to the limited information access in decentralized scenarios, STGNN may not generate strategies as effectively as the expert. Nevertheless, after being trained with historical and neighboring data and the output of the expert model, STGNN demonstrates the capability to generate strategies that approach expert-level optimization, despite the limited input data.

### A. Problem Statement

Denote the robot network at timestamp  $t$  as  $\mathcal{G}^{(t)}(\mathbf{V}^{(t)}, \mathbf{E}^{(t)}, \mathbf{X}^{(t)})$ , where  $\mathbf{V}^{(t)}$  is the set of robots,  $\mathbf{E}^{(t)}$  is the set of direct communications between nearby robots, and  $\mathbf{X}^{(t)}$  is the set of states for the robots. For an arbitrary node  $V_i$ , its neighborhood  $\mathcal{N}_i^{(t)}$  at timestamp  $t$  is composed of other robots  $V_j$  within its communication range  $R_c$ , i.e.,  $\{V_j \in \mathcal{N}_i^{(t)} | \forall V_j, r_{i,j}^{(t)} < R_c\}$ .  $r_{i,j}^{(t)}$  is the distance between  $V_i$  and  $V_j$  at  $t$ . The problem can then be summarized as below:

$$f_{\text{expert}}(X_i^{(t)}, \mathbf{V}^{(t)}, \mathbf{E}^{(t)}) \rightarrow \mathbf{u}^{(t+1)}, \quad (1)$$

$$f_{\text{STGNN}}(X_i^{(t-L)} \dots X_i^{(t)}, \mathcal{N}_i^{(t-L)} \dots \mathcal{N}_i^{(t)}, \mathbf{E}_{i,j; V_j \in \mathcal{N}_i^{(t-L)} \dots \mathcal{N}_i^{(t)}}(\theta)) \rightarrow \mathbf{u}^{(t+1)}, \quad (2)$$

$$\theta^* = \arg \min_{\theta} \left( \sum_{V_i \in \mathbf{V}} (\mathbf{u}^{(t+1)} - \mathbf{u}'^{(t+1)})^2 \right), \quad (3)$$

where  $L$  is the number of historical states used in STGNN,  $\theta$  is the set of trainable parameters of STGNN, and  $\theta^*$  is the optimized parameter set for STGNN.

### B. Expert algorithm

Previous researchers recognized the challenge of communication delays in the early stages of robot flocking studies, leading them to focus on decentralized scenarios. Consequently, only a limited number of prior algorithms have been designed to address centralized scenarios in the context of robot flocking problems. Therefore, to formulate an expert model for targeting our tasks and considering centralized scenarios, we propose a novel centralized model to generate labels for STGNN training.

As introduced in Section II, an expert algorithm should cover three crucial aspects—flocking, the leader following, and obstacle avoidance. In this case, the algorithm should be constrained by both the distances between robots and the distances between robots and obstacles. Similar to [4], the regularized update of speed is computed jointly from the collision avoidance potential and the velocity agreement. The collision avoidance potential ensures that the distances between robots exceed a predefined threshold, while the velocity agreement ensures that robots maintain consistent behavior in relation to the other robots.

Specifically, consider there exists a control  $u_i$  of robot  $i$  as the input of the expert algorithm. The algorithm then generates an update  $u_i$  following specific rules.  $u_i$  can be defined by the combination of these components as in Equation 4. In Equation 4,  $c_\alpha$ ,  $c_\beta$ , and  $c_\gamma$  are positive weighting parameters. The  $\alpha$ - term specifies collision avoidance and velocity alignment among the robots, while the  $\beta$ - term specifies collision avoidance and velocity alignment between robots to obstacles. The  $\gamma$  term is peer-to-peer guidance from the virtual leader to each robot  $i$ .

$$\mathbf{u}_i = c_\alpha \mathbf{u}_i^\alpha + c_\beta \mathbf{u}_i^\beta + c_\gamma \mathbf{u}_i^\gamma. \quad (4)$$

The collision avoidance potential  $U$  increases significantly as the distance  $r_{i,j}$  between two robots decreases. This increment is governed by a reciprocal function that dominates when  $r_{i,j}$  approaches zero (Eq. 5). Consequently,  $U_{i,j}$  captures the fact that beyond a certain distance threshold (e.g., communication range  $R_c$ ), no direct interaction exists between robots in terms of collision avoidance. The velocity agreement is the velocity difference between robot  $i$  and all the other robots (the first term in Eq. 6). The resulting control input  $u_i^\alpha$  (as defined in Eq. 6) is a centralized solution that takes into account the velocity mismatch among all robots and the local collision potential. Note that the collision avoidance implemented by Equation 5 does not guarantee a minimal distance between robots. Our results show that the minimal distance decreases when the number of robots increases.

$$U_{i,j} = \frac{1}{r_{i,j}^2} + \log \|r_{i,j}\|^2, \|r_{i,j}\| \leq R_c. \quad (5)$$

$$\mathbf{u}_i^\alpha = - \sum_{j=1}^N (\mathbf{v}_i - \mathbf{v}_j) - \sum_{j=1}^N (\nabla_{r_{i,j}} U_{i,j}). \quad (6)$$

To incorporate obstacle avoidance in the model, we follow Olfati-Saber [3]’s work by introducing an imaginary robot, i.e.,  $\beta$ -robot, which is defined by the projection of a robot  $i$  on the  $k$ -th obstacle  $O^k$  within its communication distance  $R_c$ , to assist with the obstacle avoidance task. In practice, this can be achieved by a robot, equipped with sensors, measuring the relative position and velocity between the closest point on an obstacle and itself [3]. The control input  $\mathbf{u}_i^\beta$  follows the flocking control (Eq. 6) while focusing on the potential between robot  $i$  and its projection on obstacles. Let  $\mathbf{p}_k^o$  denote the position of the  $k$ -th obstacle. Then the position and velocity of the  $\beta$ -robot, created by projecting the  $i$ -th

robot on the  $k$ -th obstacle, can be calculated by Equation 7 and Equation 8, respectively.

$$\mathbf{p}_{i,k} = \mu * \mathbf{p}_i + (1 - \mu) \mathbf{p}_k^o, \mu = \frac{r_k}{\|\mathbf{p}_i - \mathbf{p}_k^o\|}. \quad (7)$$

$$\mathbf{v}_{i,k} = \mu P \mathbf{v}_i, P = I - \mathbf{a}_k \mathbf{a}_k^T, a_k = \frac{(\mathbf{p}_i - \mathbf{p}_k^o)}{\|\mathbf{p}_i - \mathbf{p}_k^o\|}. \quad (8)$$

Define the state of the virtual leader by its position  $p^r$  and velocity  $v^r$ . The leader following control  $u^r$  can be defined in Equation 9, where  $c_1, c_2$  are positive weighting parameters.

$$\mathbf{u}_i^\gamma = -c_1(\mathbf{p}_i - \mathbf{p}^r) - c_2(\mathbf{v}_i - \mathbf{v}^r), c_1 > 0, c_2 = \sqrt{c_1}. \quad (9)$$

### C. STGNN-based imitation learning

In this section, we present an overview of the STGNN-based learning model. We begin by describing the strategy used to construct the model input. Then we discuss the architecture of the STGNN model, which involves two levels of state expansion.

a) *Local observation*: We develop a decentralized solution that only requires local observations of each robot. Inspired by Tolstaya [5], we define the local state of  $i$ -th robot in Equation 10, which consists of the aggregation from neighboring robots ( $\alpha$  term), local observation of obstacles ( $\beta$  term), and local observation of the virtual leader ( $\gamma$  term). The local observation of obstacles is defined as  $k \in \mathcal{N}$  with  $r_{i,k} < R_c$ .  $r_{i,k}$  is the distance between robot  $i$  and  $\beta$ -robot, the projection of robot  $i$  on the  $k$ -th obstacle.

$$X_i^{\alpha,\beta,\gamma} = [[\sum_{j \in \mathcal{N}} X_{i,j}^\alpha]; [\sum_{k \in \mathcal{N}} X_{i,k}^\beta]; [X_i^\gamma]]. \quad (10)$$

Instead of directly using the position  $\mathbf{p}_i$  and velocity  $\mathbf{v}_i$  of robot  $i$  as input for our model, we adopt the relative state to be consistent with the expert algorithm. We define the relative state for robot  $i$  to robot  $j$  as  $X_{i,j}$  (Eq. 11). Then we aggregate the local peer-to-peer relative states into a local state  $X_i$ , which includes  $X_i^\alpha$  (Eq. 12) and  $X_i^\beta$  (Eq. 13).

$$X_{i,j} = [\mathbf{v}_i - \mathbf{v}_j, \frac{\mathbf{p}_i - \mathbf{p}_j}{r_{i,j}}, \frac{\mathbf{p}_i - \mathbf{p}_j}{r_{i,j}^2}]. \quad (11)$$

$$X_i^\alpha = \frac{1}{N} \sum_{j \in \mathcal{N}} X_{i,j}^\alpha, j \in \mathcal{N} \text{ if } r_{i,j} < R_c. \quad (12)$$

$$X_i^\beta = \sum_{k \in \mathcal{N}} X_{i,j}^\beta, k \in \mathcal{N} \text{ if } r_{i,o} < R_c. \quad (13)$$

Notably, our model only requires local state information defined by Equation 11, which makes our approach decentralized.

b) *STGNN*: To mimic the control input  $\mathbf{u}$  generated by the expert algorithm that uses the global information, we implement two levels of expansions based on local observation as illustrated in Figure 2. The first level is spatial expansion, aiming at extracting information from additional robots, while the second level is temporal expansion, which centers on local state evolution. For each expansion, we extract information on  $L$  timestamps. For the timestamp  $t$ , consider an arbitrary node  $V_i$  with state  $X_i^{(t)}$  and 1-hop

neighborhood  $\mathcal{N}_i^{(t)}$ . A node  $V_j$  is  $V_i$ 's  $l$ -hop neighbor if it is a 1-hop neighbor of at least one  $V_i$ 's  $l-1$ -hop neighbor. Note that even though we use "l-hop neighbor" to account for state delays resulting from spatial distances, the state of one hop further is the same as one timestamp earlier in our scenario. In other words, the state of a  $l$ -hop neighbor  $j$ ,  $X_j^{(l)}$ , is equivalent to  $X_j^{(t-l+1)}$ .

In spatial expansion, we merge each node's current state  $X_i^{(t)}$  to its  $l$ -hop delayed neighbors, where  $l = 1, 2, \dots, L$ . Specifically, we first aggregate the delayed states for each hop, as indicated in Equation 14. Then,  $X_i^{(t)}$  is updated with its neighborhood information as in Equation 15.

$$X'_{i,s}{}^{(l)} = \sum_{j \in \mathcal{N}_i^{(l)}} X_j^{(l)}, l = 1, 2, \dots, L. \quad (14)$$

$$H_{i,s}^{(t)} = [X_i^{(t)} ||_{l=1, \dots, L} X'_{i,s}{}^{(l)}]. \quad (15)$$

The same method is applied to other timestamps  $l = t-L, \dots, t-1$ , but with  $t-l+1$ -hop neighbors. Then,  $H_{i,s} = \{H_{i,s}^{(l)}, l = t-L, \dots, t\}$  are fed into a transformer [15] for temporal-wise information fusion. The transformer takes three copies of  $H_{i,s}$  and considers them as query, key, and value separately. The query and key are used to compute relations between different timestamps and the value is used to generate temporal-wise fused embeddings:

$$H'_{i,s} = \sigma\left(\frac{WqH_{i,s}(WkH_{i,s})^T}{\sqrt{d}}\right)WvH_{i,s} \quad (16)$$

where  $\sigma$  is a softmax function and  $d$  is the number of features of  $H_{i,s}$ .

The temporal expansion considers the temporal evolution pattern for local spatial state fusion (i.e., considers only 1-hop neighborhood). Specifically, the spatial data fusion method is shown in Equations 14 and 15, and the temporal feature fusion is achieved by a transformer as described in Equation 16. Consider the output of the temporal expansion module as  $H'_{i,t}$ , the concatenation of  $H'_{i,s}$  and  $H'_{i,t}$  is fed to a feed-forward network to generate the final flocking controls for the robots.

$$X'_{i,t} = \sum_{j \in \mathcal{N}_i^{(1)}} X_j^{(l=1)}. \quad (17)$$

$$H_{i,t} = [X_i^{(t)} || X'_{i,t}{}^{(l=1)}]. \quad (18)$$

#### IV. SIMULATION AND EXPERIMENT

In this section, we discuss the settings of the STGNN-based learning model, evaluation metrics, simulation results, and a real robot experiment. A video of our simulation and experiment is available online.<sup>1</sup>

<sup>1</sup> <https://youtu.be/Hb70fe3IvaA>

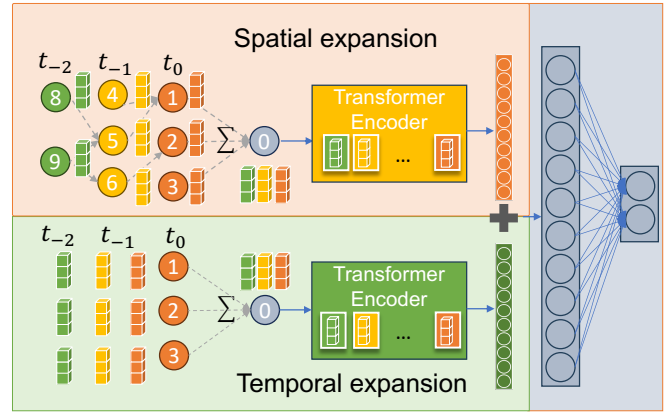


Fig. 2. STGNN spatial and temporal expansion module with  $L = 3$ . Spatial expansion and temporal expansion operate in parallel. Both spatial and temporal expanded states are concatenated and passed to the next layer of the neural network to predict the control input.

##### A. Settings for STGNN-based learning model

a) *Local observation*: The local observation (Eq. 10) denoted as  $X \in \mathbb{R}^{18}$ , is passed through a two-layer multi-layer perceptron (MLP) with a hidden size of 128 to extract the local feature embedding  $H \in \mathbb{R}^{128}$  prior to undergoing spatial and temporal expansion. Each  $X$  from the historical data follows the same process.

b) *STGNN spatial expansion*: A single SAGEConv layer with a hidden size of 128 is used to aggregate local neighbor information. Then the  $l$ -hop neighbor information is obtained using Equations 14 and 15. The result is a sequence of  $L$  delayed states. A transformer consisting of two encoder layers, each with a feedforward size of 16 and a head size of 4, is used to extract the spatial expansion results.

c) *STGNN temporal expansion*: For each local features from the history  $L$ , a single SAGEConv layer with a hidden size of 128 is used to aggregate neighbor information (Eq. 15). A transformer consisting of two encoder layers, each with a feedforward size of 16 and a head size of 4, is used to extract the temporal expansion results.

d) *Action generation*: The last output from both spatial and temporal expansions is concatenated to get the fusion embedding  $H \in \mathbb{R}^{256}$ . Subsequently, this output undergoes a two-layer MLP to generate the predicted control signal  $\hat{u}$ . The ground truth of the training data set is derived by the expert algorithm (Sec. III-B). L2 loss with the Adam optimizer is used to train the model.

We trained the STGNN model with  $N = 20$ , which includes three spherical obstacles and one virtual leader. The size of the swarm is chosen to ensure that spatial expansion could encompass a significant number of robots. The virtual leader goes through the obstacles but the robots must avoid collision with obstacles while following the virtual leader. The virtual leader moves continuously at a constant speed of 1m/s along the x-axis. The communication range is  $R_c = 1$  m, and the sampling period is  $T_s = 0.01$  s. The initial robot positions are randomized, following a uniform distribution within the range of  $[0, 0.5R_c\sqrt{N}]$ . The initial velocity is 0

m/s. The maximum allowed acceleration is  $U_{\max} = 10$  m/s<sup>2</sup> and the maximum velocity allowed is  $V_{\max} = 10$  m/s on each axis. The safety distance is 0.15 m. If  $r_{i,j}$  falls below 0.15 m, the experiment is terminated early. To ensure a valid initial configuration,  $r_{i,j}$  must be greater than 0.15 m. The training episode has 1200 steps to allow all robots to pass obstacles and form flocking behavior on the other side. One trajectory of the expert algorithm used in training is shown in Figure 3.

We implement our model using PyTorch and the OpenAI gym framework in Python3.9. The server we use has Intel(R) Xeon(R) W-2133 CPU @ 3.60GHz, NVIDIA QUADRO P5000 GPU, and 32 GB RAM. We comprehensively evaluate our proposed method, STGNN, with  $L$  set to 1, 2, 3. The compared algorithms include DGNN [5], [6], TGNN, and Olfati-Saber’s decentralized flocking algorithm denoted as Saber [3].

### B. Metrics

- 1) Completion Rate (C%): the rate of successfully completed episodes. The episode can be terminated early if any robot hits obstacles or  $r_{i,j}$  falls below 0.15 m. *The higher value is better.*
- 2) Mean Absolute Error (MAE): the mean absolute error between the expert control  $\mathbf{u}$  and model prediction  $\hat{\mathbf{u}}$ . *The lower value is better.*
- 3) Velocity alignment (V): the velocity variance of the swarm at the end of the episode. All robots should be velocity aligned at the end of the episode thus *the lower value is better.*
- 4) Distance to the leader ( $\tau$ , Eq. 19): the mean distance from any robot to the leader is an auxiliary measure. A large value indicates robots deviate from the leader and move in different directions, resulting in the failure of swarm formation. However, a small value indicates a higher risk of collision within the swarm. *Close to the expert algorithm is better.*

$$\tau = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \tau_i^t. \quad (19)$$

### C. Experiment results

We train one model for each setting described in Section IV-A. The training consists of 200 epochs with an initial learning rate of 1e-3. We implement early stopping and exponential learning rate decay to prevent overfitting.

a) *Evaluation on swarm size of 20:* In the first set of experiments, the trained models are tested in the same environment as the training environment, as described in Section IV-A. The only difference lies in the initial positions of the robots, which vary due to random initialization. During the testing phase, the swarm’s next state is determined by the model’s prediction  $\hat{\mathbf{u}}$ . If an episode is terminated early due to collision, the failed episode’s metric (MAE, V,  $\tau$ ) is not included in the aggregation for reporting. Each model runs over 20 trials and the mean and standard deviation are reported in Table I. For STGNN L1, L2, and L3 models, the

results demonstrate consistent improvement as spatial and temporal expansion increases. By increasing L from 1 to 3 in both spatial and temporal expansion, the completion rate increases from 0.85 to 1.0, the MAE decreases from 3.03 to 1.71, and the velocity variance also decreases from 0.14 to 0.01. The distance to the leader does not consistently decrease, but as explained in Section IV-B, a smaller value indicates a higher risk of collision, so both values of 1.03 and 1.06 are acceptable. When considering DGNN L3 and TGNN L3, both models exhibit improved performance compared to STGNN L1, which uses only local neighbor information. Furthermore, STGNN L3, which is the combination of DGNN L3 and TGNN L3, leads to further performance enhancement. Olfati-Saber’s decentralized solution [3] achieves a similar success rate as STGNN L1, as both models utilize only local information. However, the Saber algorithm differs from our expert algorithm, and its formations tend to have larger minimal distances, resulting in a larger  $\tau$ , which is not directly comparable to the expert algorithm.

In Figure 3, a test trajectory of the expert algorithm is plotted on the top and the test trajectory of STGNN L3 is plotted at the bottom. The plots illustrate that STGNN L3 possesses the capability to mimic the expert algorithm’s control input  $\mathbf{u}$  and generate similar flocking trajectories for the robots.

Model	C%	MAE	V	$\tau$
Expert	1.00	–	0.00 ± 0.00	0.97 ± 0.01
Saber	0.85	–	0.02 ± 0.02	2.27 ± 0.14
DGNN L3	<b>1.00</b>	2.83 ± 0.29	0.03 ± 0.11	1.08 ± 0.12
TGNN L3	<b>1.00</b>	2.11 ± 0.14	0.02 ± 0.03	<b>1.02 ± 0.05</b>
STGNN L1	0.85	3.03 ± 0.74	0.14 ± 0.34	1.32 ± 0.31
STGNN L2	0.90	2.25 ± 0.19	0.02 ± 0.03	1.03 ± 0.05
STGNN L3	<b>1.00</b>	<b>1.71 ± 0.17</b>	<b>0.01 ± 0.01</b>	1.06 ± 0.06

TABLE I

TEST RESULTS OF  $N = 20$ . L DENOTES THE HISTORY LENGTH. STGNN L3 ACHIEVES THE BEST PERFORMANCE.

b) *Evaluation on varying swarm sizes:* In the second set of experiments, we examine the model’s transferability across various swarm sizes, specifically for  $N = 30, 40,$  and  $50$ . We directly employ previously trained models to larger swarm sizes and conduct 20 trials for each scenario. The mean and standard deviation for each model and swarm size are reported in Table II. To facilitate swarm formation after obstacle avoidance, we increase the maximum episode steps from 1200 to 1500, 1800, and 2100 for swarm sizes  $N = 30, 40,$  and  $50$ , respectively.

STGNN L3 successfully achieves flocking with leader following and obstacle avoidance through all testing cases and attains the lowest MAE compared to STGNN models with shorter history. Velocity alignment is achieved in both STGNN L2 and STGNN L3. The large  $\tau$  in STGNN L1 indicates the failure in swarm formation. Our proposed expert algorithm consistently performs well in terms of success rate, velocity alignment, and distance to target, demonstrating its suitability as the ground truth. Furthermore, the performance

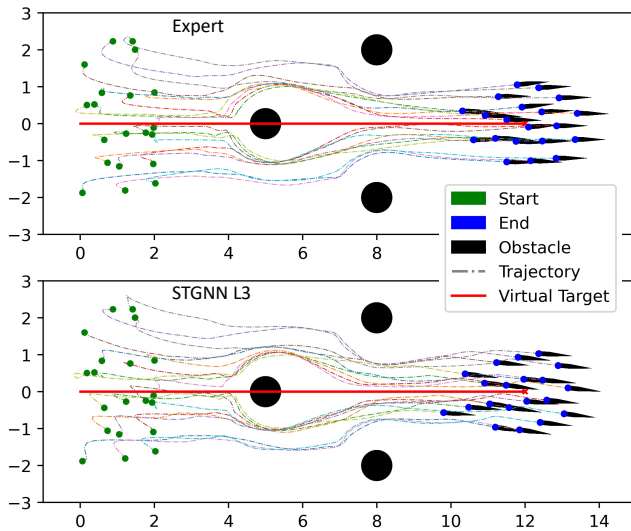


Fig. 3. Simulation result of  $N = 20$ . Top: the expert model drives 20 robots. Bottom: STGNN with  $L3$  drives 20 robots. The two models are tested in the same environment with the same initialization.

of STGNN  $L3$  underscores its ability to provide accurate estimations of the global control inputs using only local information.

Model	C%	MAE	N 30	
			V	$\tau$
Expert	1.00	–	$0.00 \pm 0.00$	$1.17 \pm 0.03$
DGNN L3	<b>1.00</b>	$3.59 \pm 0.22$	$0.11 \pm 0.20$	$1.22 \pm 0.16$
TGNN L3	0.95	$2.89 \pm 0.21$	<b>0.01</b> $\pm$ <b>0.00</b>	$1.11 \pm 0.03$
STGNN L1	0.90	$4.14 \pm 0.32$	$0.06 \pm 0.10$	$2.87 \pm 1.99$
STGNN L2	0.90	$3.19 \pm 0.17$	<b>0.01</b> $\pm$ <b>0.00</b>	$1.11 \pm 0.05$
STGNN L3	<b>1.00</b>	<b><math>2.69 \pm 0.15</math></b>	<b>0.01</b> $\pm$ <b>0.01</b>	<b><math>1.12 \pm 0.05</math></b>
Model	C%	MAE	N 40	
			V	$\tau$
Expert	1.00	–	$0.00 \pm 0.00$	$1.29 \pm 0.03$
DGNN L3	0.95	$4.54 \pm 0.20$	$0.62 \pm 0.38$	$1.73 \pm 0.28$
TGNN L3	<b>1.00</b>	$3.99 \pm 0.15$	<b>0.01</b> $\pm$ <b>0.00</b>	<b><math>1.19 \pm 0.07</math></b>
STGNN L1	0.95	$5.63 \pm 0.21$	$0.14 \pm 0.23$	$8.87 \pm 3.36$
STGNN L2	0.90	$4.43 \pm 0.11$	<b>0.01</b> $\pm$ <b>0.01</b>	$1.16 \pm 0.04$
STGNN L3	<b>1.00</b>	<b><math>3.88 \pm 0.14</math></b>	<b>0.01</b> $\pm$ <b>0.01</b>	$1.16 \pm 0.02$
Model	C%	MAE	N 50	
			V	$\tau$
Expert	1.00	–	$0.00 \pm 0.00$	$1.39 \pm 0.01$
DGNN L3	<b>1.00</b>	$5.43 \pm 0.30$	$0.89 \pm 0.77$	$2.16 \pm 0.59$
TGNN L3	<b>1.00</b>	<b><math>4.90 \pm 0.13</math></b>	<b>0.00</b> $\pm$ <b>0.00</b>	$1.23 \pm 0.09$
STGNN L1	0.95	$6.90 \pm 0.20$	$0.89 \pm 1.70$	$21.61 \pm 6.34$
STGNN L2	0.95	$5.44 \pm 0.10$	$0.01 \pm 0.01$	<b><math>1.26 \pm 0.09</math></b>
STGNN L3	<b>1.00</b>	$5.00 \pm 0.10$	<b>0.00</b> $\pm$ <b>0.00</b>	$1.23 \pm 0.11$

TABLE II

TEST RESULTS FOR  $N = 30$ ,  $N = 40$ ,  $N = 50$ . STGNN  $L3$  COMPLETES SIMULATIONS IN ALL SWARM SIZES.

#### D. Real robot experiment

We further demonstrate the effectiveness of STGNN by implementing it to achieve flocking behaviors of a group of Bitcraze Crazyflie 2.1 drones [16]. The drones are controlled using the Crazyswarm platform [17], which is based on The Robot Operating System (ROS) [18] and allows Crazyflie

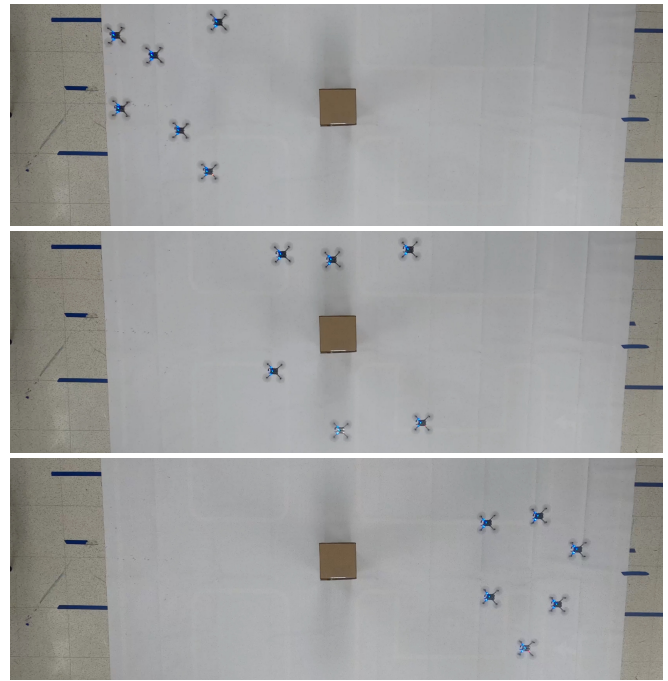


Fig. 4. The plots from top to bottom show an experiment of six drones starting from random locations, avoiding an obstacle, and achieving flocking. More experiments can be found in the video attachment.

drones to fly as a swarm in a tight and synchronized formation. The positions of the drones are obtained using the LightHouse positioning system. The system utilizes the SteamVR base stations together with the positioning deck on the drone to estimate the position of the drones. As shown in Figure 4, the six drones start with random locations (top), navigate through and avoid obstacles (middle), and form a flock on the other side of the obstacle (right). Please refer to the online video for a more detailed real robot experiment with 4, 5, and 6 drones (see footnote 1).

#### V. CONCLUSION AND FUTURE WORK

We demonstrate the effectiveness of STGNN as a decentralized solution for flocking with the leader following and obstacle avoidance tasks. STGNN overcomes the limitations of relying solely on local information by integrating prediction capabilities into the model, enabling it to capture and respond to global swarm dynamics. Our STGNN-based learning model consistently outperforms the existing decentralized algorithm introduced by Olfati-Saber [3]. Moreover, the performance of the STGNN model consistently improves as  $L$  increases as shown in Table I. Furthermore, STGNN outperforms spatial-only models, demonstrating its ability to utilize both spatial and temporal information for enhanced flocking control. The design is flexible in terms of  $L$ -hop spatial and temporal expansion, which enables seamless adaptation to various swarm sizes and history lengths. In the future, we plan to investigate the capabilities of STGNN [19]–[21] in a broader range of multi-agent tasks, including target tracking [7], [22], path planning [23], [24] and coverage and exploration [25], [26].

## REFERENCES

- [1] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," in *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, 1987, pp. 25–34.
- [2] M. Majid, M. Arshad, and R. Mokhtar, "Swarm robotics behaviors and tasks: a technical review," *Control Engineering in Robotics and Industrial Automation: Malaysian Society for Automatic Control Engineers (MACE) Technical Series 2018*, pp. 99–167, 2022.
- [3] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on automatic control*, vol. 51, no. 3, pp. 401–420, 2006.
- [4] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Stable flocking of mobile agents part i: dynamic topology," in *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475)*, vol. 2. IEEE, 2003, pp. 2016–2021.
- [5] E. Tolstaya, F. Gama, J. Paulos, G. Pappas, V. Kumar, and A. Ribeiro, "Learning decentralized controllers for robot swarms with graph neural networks," in *Conference Robot Learning 2019*. Osaka, Japan: Int. Found. Robotics Res., 30 Oct.-1 Nov. 2019.
- [6] R. Kortvelesy and A. Prorok, "Modgmn: Expert policy approximation in multi-agent systems with a modular graph neural network architecture," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 9161–9167.
- [7] L. Zhou, V. D. Sharma, Q. Li, A. Prorok, A. Ribeiro, P. Tokekar, and V. Kumar, "Graph neural networks for decentralized multi-robot target tracking," in *2022 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. IEEE, 2022, pp. 195–202.
- [8] E. S. Lee, L. Zhou, A. Ribeiro, and V. Kumar, "Graph neural networks for decentralized multi-agent perimeter defense," *Frontiers in Control Engineering*, vol. 4, p. 1104745, 2023.
- [9] C. Yan, C. Wang, X. Xiang, K. H. Low, X. Wang, X. Xu, and L. Shen, "Collision-avoiding flocking with multiple fixed-wing uavs in obstacle-cluttered environments: A task-specific curriculum-based madrl approach," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.
- [10] J. Xiao, Z. Wang, J. He, and G. Yuan, "A graph neural network based deep reinforcement learning algorithm for multi-agent leader-follower flocking," *Information Sciences*, p. 119074, 2023.
- [11] J. Orr and A. Dutta, "Multi-agent deep reinforcement learning for multi-robot applications: a survey," *Sensors*, vol. 23, no. 7, p. 3625, 2023.
- [12] G. Swamy, S. Choudhury, D. Bagnell, and S. Wu, "Sequence model imitation learning with unobserved contexts," in *Advances in Neural Information Processing Systems*, A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho, Eds., 2022. [Online]. Available: <https://openreview.net/forum?id=3nbKUphLBg5>
- [13] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in neural information processing systems*, vol. 30, 2017.
- [14] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.
- [15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [16] "Bitcraze crazyflie," <https://www.bitcraze.io/>, accessed: 2023-09-15.
- [17] J. A. Preiss\*, W. Hönig\*, G. S. Sukhatme, and N. Ayanian, "Crazyswarm: A large nano-quadcopter swarm," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 3299–3304, software available at <https://github.com/USC-ACTLab/crazyswarm>. [Online]. Available: <https://doi.org/10.1109/ICRA.2017.7989376>
- [18] "Robot operating system," <https://www.ros.org/>, accessed: 2023-09-15.
- [19] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph wavenet for deep spatial-temporal graph modeling," *arXiv preprint arXiv:1906.00121*, 2019.
- [20] M. Xu, W. Dai, C. Liu, X. Gao, W. Lin, G.-J. Qi, and H. Xiong, "Spatial-temporal transformer networks for traffic flow forecasting," *arXiv preprint arXiv:2001.02908*, 2020.
- [21] S. Hadou, C. I. Kanatsoulis, and A. Ribeiro, "Space-time graph neural networks with stochastic graph perturbations," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.
- [22] L. Zhou, V. Tzoumas, G. J. Pappas, and P. Tokekar, "Resilient active target tracking with multiple robots," *IEEE Robotics and Automation Letters*, vol. 4, pp. 129–136, 2018. [Online]. Available: <https://api.semanticscholar.org/CorpusID:52189660>
- [23] Q. Li, F. Gama, A. Ribeiro, and A. Prorok, "Graph neural networks for decentralized multi-robot path planning," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 11 785–11 792.
- [24] Q. Li, W. Lin, Z. Liu, and A. Prorok, "Message-aware graph attention networks for large-scale multi-robot path planning," *IEEE Robotics and Automation Letters*, vol. 6, pp. 5533–5540, 2020. [Online]. Available: <https://api.semanticscholar.org/CorpusID:227209291>
- [25] E. Tolstaya, J. Paulos, V. Kumar, and A. Ribeiro, "Multi-robot coverage and exploration using spatial graph neural networks," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 8944–8950.
- [26] V. D. Sharma, L. Zhou, and P. Tokekar, "D2coplan: A differentiable decentralized planner for multi-robot coverage," *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3425–3431, 2022. [Online]. Available: <https://api.semanticscholar.org/CorpusID:252383510>