

# Distributed Resilience-Aware Control in Multi-Robot Networks

Haejoon Lee<sup>1</sup> and Dimitra Panagou<sup>1,2</sup>

**Abstract**—Ensuring resilient consensus in multi-robot systems with misbehaving agents remains a challenge, as many existing network resilience properties are inherently combinatorial and globally defined. While previous works have proposed control laws to enhance or preserve resilience in multi-robot networks, they often assume a fixed topology with known resilience properties, or require global state knowledge. These assumptions may be impractical in physically-constrained environments, where safety and resilience requirements are conflicting, or when misbehaving agents share inaccurate state information. In this work, we propose a distributed control law that enables each robot to guarantee resilient consensus and safety during its navigation without fixed topologies using only locally available information. To this end, we establish a sufficient condition for resilient consensus in time-varying networks based on the degree of non-misbehaving or normal agents. Using this condition, we design a Control Barrier Function (CBF)-based controller that guarantees resilient consensus and collision avoidance without requiring estimates of global state and/or control actions of all other robots. Finally, we validate our method through simulations.

## I. INTRODUCTION

The consensus problem in multi-agent systems finds numerous applications from control to optimization [1], [2]. However, the performance of consensus in general degrades in the presence of misbehaving agents that share faulty or incorrect information. Resilient consensus has therefore been extensively studied [3]–[7]. In particular, the *Weighted Mean-Subsequence-Reduced* (W-MSR) algorithm was introduced in [3] to guarantee consensus among non-misbehaving or normal agents despite the presence of misbehaving agents.

The challenge of using W-MSR algorithm lies in its reliance on complex global topological network resilience properties, such as  $r$ -robustness and  $(r, s)$ -robustness [3]. These properties are inherently combinatorial, making them difficult to compute online [8]. Moreover, in many multi-robot applications, links between robots are formed if their relative distances are within a threshold, which further complicates the practical implementation of these network properties in real-time operation.

To implement these network properties, [9], [10] enforce local inter-robot connectivity by maintaining predetermined topologies with known resilience levels. This approach eliminates the need for continuous resilience computation. However, its limitation is that the imposed connectivity constraints induce fixed topologies, making it impractical

in dynamic or constrained environments where robots need flexible movements. There is earlier work considering the maintenance of network resilience without relying on fixed topologies [9], [11]–[15], but these resilience-aware approaches rely on global state knowledge (i.e., states of all robots). To obtain global state knowledge, each robot needs to accurately estimate it based on the information from its neighbor robots, which becomes challenging as information being shared is unreliable in adversarial settings.

In this paper, we focus on the design of distributed controllers based on Control Barrier Functions (CBFs) [16], where each robot computes its control inputs using locally-available information only, so that it maintains a resilient structure and ensures resilient consensus without fixed topologies. The problem of developing CBF-based controllers for improving or maintaining resilience of multi-robot networks without relying on fixed network topologies has been studied in [11], [13]–[15]. However, these approaches take a centralized CBF form, or assume that each robot can accurately estimate the global state knowledge and the control actions of all other robots at every time instant. These assumptions are unrealistic in complex and adversarial networks, and difficult to resolve due to the global nature of the resilience properties.

To address these problems, the main contribution of this paper is twofold. We develop a sufficient condition for resilient consensus in time-varying networks based only on the degree (i.e., the number of neighbors) of normal agents. While [7] establishes similar conditions based on degree of all agents in continuous-time consensus dynamics, our condition applies to discrete-time dynamics and specifically to the W-MSR algorithm. Then, leveraging this condition, we design a distributed controller that maintains each normal robot's degree above a threshold while preventing collisions. We prove that our controller guarantees resilience and safety using only local information under certain assumptions. In contrast to earlier work [11]–[14], our approach (a) focuses on local connectivity of normal agents rather than enforcing global resilience properties, (b) eliminates reliance on potentially unreliable global state estimation, and (c) explicitly accounts for inter-agent collisions, which is overlooked in previous CBF-based approaches [11], [13], [14], [17]–[19].

**Organization:** In Section II, we introduce the notation, preliminaries, and the problem statement. In Section III we present a sufficient condition for resilient consensus, and in Section IV we design a distributed CBF-based controller for resilience. Simulation results evaluate our approach in Section V, while Section VI concludes our work.

\*This work is supported by the Air Force Office of Scientific Research (AFOSR) under FA9550-23-1-0163

<sup>1</sup>Department of Robotics, University of Michigan, Ann Arbor, MI, USA  
haejoonl@umich.edu

<sup>2</sup>Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA  
dpanagou@umich.edu

## II. PRELIMINARIES AND PROBLEM STATEMENT

We denote the cardinality of a set  $\mathcal{S}$  as  $|\mathcal{S}|$ . We use  $\mathbb{R}_+$ ,  $\mathbb{R}_{\geq 0}$ , and  $\mathbb{Z}_{\geq 0}$  to denote sets of positive real, non-negative real, and non-negative integer numbers, respectively. We also use  $\|\cdot\| : \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}$  to represent the Euclidean norm.

### A. Multi-Robot Network System

Consider a multi-robot system of  $n$  robots. Robot  $i \in \{1, \dots, n\}$  has a physical state  $x_i(t) \in \mathbb{R}^m$  and control input  $u_i(t) \in \mathcal{U}_i \subset \mathbb{R}^m$ , denoting position and velocity at time  $t$ , respectively. Each robot  $i$  has single integrator dynamics as

$$\dot{x}_i(t) = u_i(t). \quad (1)$$

We model robots as circular disks with radii  $\frac{\Delta_d}{2}$ , where  $\Delta_d \in \mathbb{R}_+$ . We also define the state and control input of the entire system as  $\mathbf{x}(t) = [x_1(t)^T \dots x_n(t)^T]^T \in \mathbb{R}^{nm}$  and  $\mathbf{u}(t) = [u_1(t)^T \dots u_n(t)^T]^T \in \mathcal{U} = \mathcal{U}_1 \times \dots \times \mathcal{U}_n \subset \mathbb{R}^{nm}$ . Then the system dynamics is

$$\dot{\mathbf{x}}(t) = \mathbf{u}(t). \quad (2)$$

The distance between robots  $i$  and  $j$  is denoted as  $\Delta_{ij}(\mathbf{x}(t)) = \|x_i(t) - x_j(t)\|$ . We model the multi-robot network as a time-varying simple, undirected graph  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ , where the vertex set  $\mathcal{V} = \{1, \dots, n\}$  includes robots, and the edge set  $\mathcal{E}(t) = \{(i, j) \mid \Delta_{ij}(\mathbf{x}(t)) \leq R\}$  defines connections (links) between two robots within a communication radius  $R \in \mathbb{R}_+$ . The set of neighbors of robot  $i \in \mathcal{V}$  at time  $t$  is denoted as  $\mathcal{N}_i(t) = \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}(t)\}$ . We also denote  $\mathcal{B}_i(t) = \mathcal{N}_i(t) \cup \{i\}$ . Each robot can observe states of its own and of its neighbors. The degree of robot  $i$  at time  $t$  is denoted as  $\delta_i(\mathcal{G}(t)) = |\mathcal{N}_i(t)|$ . Then, the minimum degree of  $\mathcal{G}(t)$  is denoted as  $\delta_{\min}(\mathcal{G}(t)) = \min_{i \in \mathcal{V}} \delta_i(\mathcal{G}(t))$ .

We smoothly approximate the network connections through the weighted adjacency matrix  $\mathbf{A}(\mathbf{x}(t))$ , whose entries are

$$a_{ij}(\mathbf{x}(t)) = \begin{cases} \frac{q_1}{1+e^{-q_2(R^2 - \Delta_{ij}(\mathbf{x}(t))^2)}} - \frac{q_1}{2} & \Delta_{ij}(\mathbf{x}(t)) \leq R, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where  $q_1, q_2 \in \mathbb{R}_+$ . For simplicity, we drop  $t$  from  $\mathbf{x}(t)$ ,  $x_i(t)$ ,  $\mathbf{u}(t)$ , and  $u_i(t)$  when the context is clear. The function (3) is continuously differentiable with respect to  $\mathbf{x}$ . At discrete times  $t_l = l\tau_1$  for  $l \in \mathbb{Z}_{\geq 0}$ , where  $\tau_1 \in \mathbb{R}_+$  is an update interval, robot  $i$  shares its connectivity level  $c_i(t_l) \in \mathbb{R}_{\geq 0}$  with its neighbors. Robot  $i$  maintains and updates its connectivity level as

$$c_i(t) = \sum_{j \in \mathcal{N}_i(t_l)} a_{ij}(\mathbf{x}(t_l)), \quad \forall t \in [t_l, t_{l+1}). \quad (4)$$

Each neighbor who receives it also maintains the value of (4) for  $t \in [t_l, t_{l+1})$  until the updated value is shared at  $t_{l+1}$ . To simplify our analysis on our main results, we assume  $\tau_1$  is sufficiently small (i.e.,  $\tau_1 \rightarrow 0$ ) such that

$$c_i(t) = \sum_{j \in \mathcal{N}_i(t)} a_{ij}(\mathbf{x}(t)) \quad (5)$$

for all  $t$ ; that is, each agent shares with neighbors in real time.

### B. Weighted Mean Subsequence Reduced Algorithm

At discrete times  $t_p = p\tau_2$  for  $p \in \mathbb{Z}_{\geq 0}$  where  $\tau_2 \in \mathbb{R}_+$  (not necessarily the same as  $\tau_1$ ) is the update interval, robot  $i$  also shares its consensus state  $y_i(t_p) \in \mathbb{R}$  with its neighbors. Robot  $i$  has  $y_i(t) = y_i(t_p) \forall t \in [t_p, t_{p+1})$  and an update protocol:

$$y_i(t_{p+1}) = \sum_{j \in \mathcal{B}_i(t_p)} w_{ij}(t_p) y_j(t_p), \quad (6)$$

where  $w_{ij}(t_p)$  is the weight assigned to  $y_j(t_p)$  by agent  $i$ . We assume  $\exists \beta \in (0, 1)$  such that  $\forall i \in \mathcal{V} \forall p \in \mathbb{Z}_{\geq 0}$ ,

- $w_{ij}(t_p) \geq \beta$  if  $j \in \mathcal{B}_i(t_p)$ , or  $w_{ij}(t_p) = 0$  otherwise,
- $\sum_{j=1}^n w_{ij}(t_p) = 1$ .

Agents are guaranteed to achieve consensus on  $y_i(t)$  through the protocol (6) if there exists a uniformly bounded sequence of contiguous time intervals where the union of  $\mathcal{G}(t)$  in each interval is connected [20, Theorem 3.10]. However, when  $\mathcal{G}$  has misbehaving and non-misbehaving (normal) agents, the consensus guarantee no longer holds. In this paper, we consider one specific type of misbehaving agents:

**Definition 1 (malicious robot).** A robot  $i \in \mathcal{V}$  is **malicious** if it sends  $c_i(T_1)$  and  $y_i(T_2)$  to all its neighbors at times  $T_1$  and  $T_2$ , respectively, but does not follow the nominal update protocols (5) and/or (6) at some times  $T_1$  and/or  $T_2$ .

We define sets of normal and malicious agents as  $\mathcal{L} \subseteq \mathcal{V}$  and  $\mathcal{A} \subseteq \mathcal{V}$ , respectively, such that  $\mathcal{L} \cap \mathcal{A} = \emptyset$  and  $\mathcal{L} \cup \mathcal{A} = \mathcal{V}$ . Various threat models describe the number of malicious robots in a network [3], and we consider the following model:

**Definition 2 (F-total).** A set  $\mathcal{S} \subset \mathcal{V}$  is **F-total** if it contains at most  $F$  nodes in the graph (i.e.  $|\mathcal{S}| \leq F$ ).

Then resilient consensus is defined as [3]:

**Definition 3 (resilient consensus).** Under an  $F$ -total attack, the normal nodes are said to achieve **resilient consensus** if both of the conditions hold for any initial consensus states:

- $\exists M \in \mathbb{R}$  such that  $\lim_{p \rightarrow \infty} y_i(t_p) = M \forall i \in \mathcal{L}$ .
- $y_i(t_p) \in [\underline{z}(0), \bar{z}(0)]$ ,  $\forall i \in \mathcal{L}$ ,  $\forall p \in \mathbb{Z}_{\geq 0}$ , where  $\underline{z}(0) = \min_{i \in \mathcal{L}} y_i(0)$  and  $\bar{z}(0) = \max_{i \in \mathcal{L}} y_i(0)$ .

**Remark 1.** Robots share two different values,  $c_i(t)$  and  $y_i(t)$ , at different update intervals. Notice that while malicious robots may manipulate one or both values by not following the nominal protocols (5) and/or (6), normal robots focus solely on achieving resilient consensus on  $y_i(t)$ .

In [3], the W-MSR (Weighted-Mean Subsequent Reduced) algorithm was introduced to guarantee resilient consensus. A normal robot  $i \in \mathcal{L}$  performs W-MSR algorithm with parameter  $F$  at every time step  $t_p$  as follows:

- 1) Robot  $i$  collects values  $y_j(t_p)$  from its neighbors  $j \in \mathcal{N}_i(t_p)$  and sorts them.
- 2) If there are at least  $F$  strictly larger values than its own  $y_i(t_p)$ , it discards the largest  $F$ ; otherwise, it removes

all larger values. Similarly, it does the same for values less than  $y_i(t_p)$ .

- 3) Let  $\mathcal{R}_i(t_p)$  be the set of robots whose values are discarded in Step 2. Then, it applies

$$y_i(t_{p+1}) = \sum_{j \in \mathcal{B}_i(t_p) \setminus \mathcal{R}_i(t_p)} w_{ij}(t_p) y_j(t_p), \quad (7)$$

where  $w_{ij}(t_p)$  satisfies the same conditions stated for the nominal consensus update protocol (6).

W-MSR algorithm ensures that normal robots can achieve resilient consensus despite the presence of up to  $F$  malicious agents, under certain topological properties of the communication graph. We review these relevant properties below:

**Definition 4 ((r,s)-reachable).** Given a graph  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$  and a nonempty subset  $\mathcal{S} \subset \mathcal{V}$ ,  $\mathcal{S}$  is **(r,s)-reachable** at time  $t$  if there are at least  $s \in \mathbb{Z}_{\geq 0}$  nodes in  $\mathcal{S}$  with at least  $r \in \mathbb{Z}_{\geq 0}$  neighbors outside of  $\mathcal{S}$  at  $t$ ; i.e., given  $\mathcal{X}_{\mathcal{S}}^r(t) = \{i \in \mathcal{S} : |\mathcal{N}_i(t) \setminus \mathcal{S}| \geq r\}$ ,  $|\mathcal{X}_{\mathcal{S}}^r(t)| \geq s$ .

**Definition 5 ((r,s)-robust).** A graph  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$  is **(r,s)-robust** at time  $t$  if for every pair of nonempty, disjoint subsets  $\mathcal{S}_1, \mathcal{S}_2 \subset \mathcal{V}$ , at least one of the following holds at  $t$  (recall  $\mathcal{X}_{\mathcal{S}_k}^r(t) = \{i \in \mathcal{S}_k : |\mathcal{N}_i(t) \setminus \mathcal{S}_k| \geq r\}$  for  $k \in \{1, 2\}$ ):

- $|\mathcal{X}_{\mathcal{S}_1}^r(t)| = |\mathcal{S}_1|$ ;
- $|\mathcal{X}_{\mathcal{S}_2}^r(t)| = |\mathcal{S}_2|$ ;
- $|\mathcal{X}_{\mathcal{S}_1}^r(t)| + |\mathcal{X}_{\mathcal{S}_2}^r(t)| \geq s$ .

A network being  $(F+1, F+1)$ -robust at all times is a sufficient condition for its normal robots to achieve resilient consensus through the W-MSR algorithm under an  $F$ -total malicious attack [3, Corollary 1]. Note for a graph of  $n$  nodes,  $F \leq \lfloor \frac{n-1}{2} \rfloor$ , which is a direct application of [3, Lemma 4]. Computing  $(r,s)$ -robustness of a network is challenging due to its combinatorial nature [8], making its integration into multi-robot systems difficult. We build on results from this notion of resilience to present an alternative sufficient condition based on the minimum degree of normal agents, offering a more computationally efficient approach.

### C. Control Barrier Function

In this work, we leverage Control Barrier Functions (CBFs) [16] to ensure resilient consensus despite an  $F$ -total attack. Note we are considering time-invariant CBFs. Let a continuously differentiable function  $h : \mathcal{D} \subset \mathbb{R}^{nm} \rightarrow \mathbb{R}$  define a safe set  $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^{nm} \mid h(\mathbf{x}) \geq 0\}$  and its boundary  $\partial\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^{nm} \mid h(\mathbf{x}) = 0\}$ . We say  $h$  is a CBF for the system (2) if there exists an extended class  $\mathcal{K}_{\infty}$  function  $\alpha : \mathbb{R} \rightarrow \mathbb{R}$  such that

$$\sup_{\mathbf{u} \in \mathcal{U}} \left[ \frac{\partial h}{\partial \mathbf{x}} \dot{\mathbf{x}} + \alpha(h(\mathbf{x})) \right] \geq 0, \quad \forall \mathbf{x} \in \mathcal{C}. \quad (8)$$

If  $h$  is a CBF on a given set  $\mathcal{C}$  for given system dynamics and  $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}) \neq 0 \forall \mathbf{x} \in \partial\mathcal{C}$ , then any locally Lipschitz continuous controller  $\mathbf{u}^*(\mathbf{x}) \in K_{\text{cbf}}(\mathbf{x}) = \{\mathbf{u} \in \mathcal{U} \mid \frac{\partial h}{\partial \mathbf{x}} \dot{\mathbf{x}} \geq -\alpha(h(\mathbf{x}))\}$  renders the set  $\mathcal{C}$  forward invariant [16, Theorem 2].

### D. Problem Statement

Let  $\mathcal{G}(t)$  be a multi-robot network formed by  $n$  robots' physical states  $\mathbf{x}$  at time  $t \geq t_0$ , where  $t_0 \in \mathbb{R}_{\geq 0}$  is the initial time. Each robot  $i \in \mathcal{V}$  also has its own connectivity level  $c_i(t)$  and consensus state  $y_i(t)$ , which it shares with its neighbors. Robots are tasked with achieving consensus on  $y_i(t)$  while navigating using a desired controller  $u_{i,\text{des}} : \mathbb{R}^m \rightarrow \mathcal{U}_i$ . Our objective is to design distributed control strategies that enable normal robots to achieve resilient consensus without fixed communication topologies, while also minimally deviating from their desired controllers. Ensuring resilient consensus in this system requires control laws for both the consensus states  $y_i(t)$  and the physical states  $x_i(t)$  of robots, which we address separately. This paper focuses on cases where malicious robots disrupt consensus through information only, that is:

**Assumption 1.** Malicious robots follow the same control laws for their physical states as normal robots and do not engage in physical disruptions (e.g., collisions).

This assumption is reasonable from the perspective that, in principle, normal robots may have detection mechanisms to detect anomalous behaviors; hence it is reasonable to assume that malicious robots aim to mimic the behavior of the normal agents' physical states to remain undetected. The consideration of both physical and informational attacks is future work.

To update consensus states, robots adopt W-MSR algorithm [3], which ensures resilient consensus given that the network topology is  $(F+1, F+1)$ -robust  $\forall t \geq t_0$ . The main challenge of using W-MSR algorithm lies in the fact that  $(r,s)$ -robustness is combinatorial and a function of global network states (i.e., the states of all robots). Existing approaches for maintaining these properties typically require obtaining global state information through inter-agent communication. However, such communication becomes unreliable in the presence of malicious agents. Thus, we present an alternative sufficient condition that is locally controllable.

**Problem 1.** Given a network  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$  under an  $F$ -total attack at time  $t$ , let  $\delta_{\mathcal{L}}(\mathcal{G}(t)) = \min_{i \in \mathcal{L}}(\delta_i(\mathcal{G}(t)))$  be the minimum degree of normal agents at time  $t$ . We aim to find the lower bound for  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \forall t \geq t_0$  in which resilient consensus is guaranteed.

Addressing Problem 1 reduces the challenge of ensuring resilient consensus with  $F$  malicious agents to ensuring each normal robot maintains its number of neighbors above a threshold  $F' \in \mathbb{Z}_+$ . The value  $F'$ , which depends on  $F$ , is provided in Proposition 1. Given Assumption 1, we solve:

**Problem 2.** Let  $\mathcal{G}(t_0) = (\mathcal{V}, \mathcal{E}(t_0))$  be a network under an  $F$ -total attack, where  $\delta_{\mathcal{L}}(\mathcal{G}(t_0)) \geq F'$ . Let  $u_{i,\text{des}}$  be the desired controller for agent  $i \in \mathcal{V}$ . We design a distributed control strategy for all robots that requires only locally available information to ensure  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \geq F'$  without collisions, while minimally deviating from  $u_{i,\text{des}}$ ,  $\forall t \geq t_0$ .

We propose a control strategy (given in (20)) such that



when all agents implement it using local information, they together ensure that normal robots maintain at least  $F'$  neighbors without collisions under some assumptions, thereby achieving resilient consensus.

### III. DEGREE AND RESILIENT CONSENSUS

Consider a graph  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$  formed by  $n$  robots' states  $\mathbf{x}$  at time  $t \in \mathbb{R}_{\geq 0}$ . In this section, we solve Problem 1 by showing that having  $\delta_{\mathcal{L}}(\mathcal{G}(t)) = \min_{i \in \mathcal{L}}(\delta_i(\mathcal{G}(t)))$  above some given threshold  $F' \in \mathbb{Z}_+$   $\forall t \geq t_0$  guarantees resilient consensus. The main advantage of using  $\delta_{\mathcal{L}}(\mathcal{G}(t))$  is that normal robots can directly control and observe their influence on it as they navigate at time  $t$ . The condition is given in Proposition 1, but before that, we provide a useful result from [21, Property 5.23] that characterizes the  $(r, s)$ -robustness of  $\mathcal{G}(t)$  in terms of its minimum degree.

**Lemma 1.** *Let  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$  be a graph with the minimum degree of  $\delta_{\min}(\mathcal{G}(t))$  at time  $t$ . Then, if  $\delta_{\min}(\mathcal{G}(t)) \geq \lfloor \frac{n}{2} \rfloor - 1$ ,  $\mathcal{G}(t)$  is  $(\delta_{\min}(\mathcal{G}(t)) - \lfloor \frac{n}{2} \rfloor + 1, s)$ -robust, for all  $1 \leq s \leq n$  [21].*

Lemma 1 provides a sufficient condition for  $(r, s)$ -robustness of a graph  $\mathcal{G}(t)$  using its minimum degree  $\delta_{\min}(\mathcal{G}(t))$  at time  $t$ . In the context of Lemma 1, the guarantees of  $(r, s)$ -robustness and hence resilient consensus depend on the local connectivity of all nodes, both normal and malicious agents. We relax this dependence to derive a sufficient condition for resilient consensus based solely on the connectivity of the *normal* agents.

**Proposition 1.** *Let  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$  be a time-varying graph under an  $F$ -total malicious attack. If  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \geq F' = F + \lfloor \frac{n}{2} \rfloor$   $\forall t \geq t_0$ , normal agents are guaranteed to achieve resilient consensus through the W-MSR algorithm with parameter  $F$ .*

*Proof.* Let  $\mathcal{G}_{\mathcal{L}}(t) = (\mathcal{L}, \mathcal{E}_{\mathcal{L}}(t))$  be a subgraph of  $\mathcal{G}(t)$  induced by the set of normal agents  $\mathcal{L}$  with  $n' := |\mathcal{L}| = n - F \geq F + 1$ . First, we analyze  $\mathcal{G}_{\mathcal{L}}(t)$ 's connectivity. All  $n'$  normal agents must have at least  $\lfloor \frac{n'}{2} \rfloor = \lfloor \frac{n'+F}{2} \rfloor$  normal neighbors. Then by [22, 3.8],  $\mathcal{G}_{\mathcal{L}}(t)$  is connected  $\forall t \geq t_0$ .

Now, consider two time-varying graphs of  $n$  nodes: (a)  $\mathcal{G}_1(t) = (\mathcal{V}_1, \mathcal{E}_1(t))$  such that  $\delta_{\min}(\mathcal{G}_1(t)) \geq F' \forall t \geq t_0$  and (b)  $\mathcal{G}_2(t) = (\mathcal{V}_2, \mathcal{E}_2(t))$  such that  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \geq F' \forall t \geq t_0$  and  $\exists T \geq t_0$  such that  $\delta_{\min}(\mathcal{G}_2(T)) < F'$ . Since  $\delta_{\min}(\mathcal{G}_1(t)) \geq F' \forall t \geq t_0$ , all agents (both normal and malicious agents) in  $\mathcal{G}_1(t)$  have at least  $F'$  neighbors  $\forall t \geq t_0$ . On the contrary, all normal agents in  $\mathcal{G}_2(t)$  have at least  $F'$  neighbors, whereas at least one malicious agent  $i \in \mathcal{A}$  in  $\mathcal{G}_2(t)$  has less than  $F'$  neighbors at  $t = T$ . By Lemma 1,  $\mathcal{G}_1(t)$  is  $(F + 1, F + 1)$ -robust for all  $t$ , which is a sufficient condition for the achievement of resilient consensus under  $F$ -total malicious attacks [3, Corollary 1]. Thus in  $\mathcal{G}_1(t)$ , normal agents have enough redundancy in information from normal agents by having  $F'$  neighbors  $\forall t \geq t_0$  to achieve resilient consensus through the W-MSR algorithm. Now, assume to the contrary that resilient consensus is not achieved among normal agents in  $\mathcal{G}_2(t)$ . Note that resilient consensus is achieved when all

normal agents achieve consensus, regardless of the consensus states of malicious agents. Furthermore, in  $\mathcal{G}_2(t)$ , normal agents have  $F'$  or more neighbors  $\forall t \geq t_0$ . This implies that they also have enough redundancy in information from normal agents  $\forall t \geq t_0$  to achieve resilient consensus through W-MSR algorithm, which is a contradiction.  $\square$

Note that the same bound for the entire graph (including both normal and malicious agents) appears in [7, Corollary 1], but in the context of continuous-time consensus dynamics and a different algorithm. Based on Proposition 1, to guarantee resilient consensus under  $F$ -total malicious attacks, it suffices that robots navigate the environment such that *normal robots* have at least  $F' = F + \lfloor \frac{n}{2} \rfloor$  neighbors  $\forall t \geq t_0$ . Since the numbers of total robots  $n$  and malicious robots  $F$  are fixed, each robot  $i \in \mathcal{V}$  can compute  $F'$  locally. Now we design a control law for the physical states of robots so that normal robots always maintain  $F'$  neighbors to achieve resilient consensus without collisions.

### IV. DISTRIBUTED RESILIENCE-AWARE CONTROLLER

In this section, we design a distributed controller for robots' physical states that enables robots to navigate the environment with guarantees of resilient consensus and physical safety. Consider a multi-robot network  $\mathcal{G}(t)$  formed by the physical states  $\mathbf{x}(t)$  of  $n$  robots at time  $t \geq t_0$ . By Proposition 1,  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \geq F' = F + \lfloor \frac{n}{2} \rfloor$   $\forall t \geq t_0$  guarantees resilient consensus through the W-MSR algorithm. Thus, we design a controller that ensures that all normal robots maintain  $F'$  or more neighbors without collisions  $\forall t \geq t_0$ , thereby addressing Problem 2.

There exist many works on decentralized control for connectivity maintenance [17], [18], [23]. However, these methods either (a) require robots to obtain states of all robots through inter-agent communication, which becomes unreliable in the presence of malicious agents, or (b) enforce fixed topologies, which may not always be feasible in certain settings or environments. Thus, we design a controller that (a) uses only local information while being robust to malicious sharing and (b) preserves a flexible network formation.

How we construct the controller (given in (20)) is as follows: first, we define constraint functions to enforce degree maintenance and collision avoidance. These functions are then composed into one centralized candidate CBF (11) whose corresponding CBF constraint (12) involves global state knowledge and control actions of all robots. Next, we decompose the constraint to create  $n$  separate *distributed* constraints (15) - one constraint for each robot - that only requires local information. Finally, we construct the controller with these distributed constraints, proving that it guarantees degree maintenance and safety under some assumptions.

#### A. Constraint with Global State Knowledge

1) *Constraint for Degree Maintenance:* We define the function that encodes maintenance of  $F'$  neighbors for robot  $i \in \mathcal{V}$ :

$$h_i(\mathbf{x}) = \sum_{j \in \mathcal{N}_i(t)} a_{ij}(\mathbf{x}) - F'. \quad (9)$$

Since  $a_{ij}(\mathbf{x})$  (3) is continuously differentiable,  $h_i(\mathbf{x})$  is also continuously differentiable. Note that although we define  $n$  CBFs so that all robots (both normal and malicious) have the same controller by Assumption 1, as will be shown in Remark 4, the guarantee of maintaining  $F'$  neighbors may apply only to normal agents.

**Lemma 2.** Let  $a_{ij}(\mathbf{x})$  be computed from (3) with  $q_1 = 2 + \epsilon$ , where  $0 < \epsilon < \frac{1}{n-1}$ , and  $q_2 \in \mathbb{R}_+$  at time  $t$ . Then,  $h_i(\mathbf{x}) \geq 0$  only if  $\delta_i(\mathcal{G}(t)) \geq F'$  at time  $t$ .

*Proof.* With  $q_1 = 2 + \epsilon$ , where  $0 < \epsilon < \frac{1}{n-1}$ , and  $q_2 \in \mathbb{R}_+$  at time  $t$ ,  $0 \leq a_{ij}(\mathbf{x}) < 1 + \frac{1}{n-1}$ . Then, given  $K = \delta_i(\mathcal{G}(t)) \leq n-1$ ,  $\sum_{j \in \mathcal{N}_i(t)} a_{ij}(\mathbf{x}) < K + \frac{K}{n-1} \leq K+1 \forall t \geq t_0$ . Thus,  $\sum_{j \in \mathcal{N}_i(t)} a_{ij}(\mathbf{x}) \geq F'$  only if  $\delta_i(\mathcal{G}(t)) \geq F'$ .  $\square$

With Lemma 2,  $h_i(\mathbf{x}) \geq 0$  only if robot  $i$  has at least  $F'$  neighbors. Thus, if  $\mathbf{x} \in \mathcal{C}_i = \{\mathbf{x} \in \mathbb{R}^{nm} \mid h_i(\mathbf{x}) \geq 0\}$  at time  $t$ , all robots are positioned in such a way that  $\delta_i(\mathcal{G}(t)) \geq F'$  at  $t$ . Since we want  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \geq F' \forall t \geq t_0$ , the safe set of interest is  $\mathcal{C} = \cap_{i \in \mathcal{L}} \mathcal{C}_i$ . Also, (9) does not specify which neighbors each robot should connect to, allowing robots the flexibility to form links dynamically. This enables the overall network to be time-varying without fixed topologies.

2) *Constraint for Inter-Agent Collision Avoidance:* For realistic deployment, we also incorporate collision avoidance between agents to ensure safe navigation. In this work, we construct candidate CBFs similar to those in [13], [19] to enforce collision avoidance between agents  $i$  and  $j$ :

$$h_{ij}^{\text{col}}(\mathbf{x}) = \Delta_{ij}(\mathbf{x})^2 - \Delta_d^2. \quad (10)$$

When  $h_{ij}^{\text{col}}(\mathbf{x}) \geq 0$ , this implies robots  $i$  and  $j$  are at least  $\Delta_d$  away from each other, ensured to be free of collision.

3) *Composition of Multiple Constraints:* We compose all constraint functions into one candidate CBF. Composition methods are studied in [24], [25], and in this paper we use the form inspired by [24]. Then, we define:

$$\phi(\mathbf{x}, \mathbf{w}) = 1 - \sum_{i \in \mathcal{V}} E_i(\mathbf{x}) - \sum_{(i,j) \in \mathcal{E}(t)} E_{ij}^c(\mathbf{x}), \quad (11)$$

where  $E_i(\mathbf{x}) = e^{-w_r h_i(\mathbf{x})}$  and  $E_{ij}^c(\mathbf{x}) = e^{-w_c h_{ij}^{\text{col}}(\mathbf{x})}$  with a parameter  $\mathbf{w} = [w_r, w_c] \in \mathbb{R}_+^2$ . Its structure implies that  $\phi(\mathbf{x}, \mathbf{w}) \geq 0$  only if  $h_i(\mathbf{x}) > 0$  and  $h_{ij}^{\text{col}}(\mathbf{x}) > 0 \forall j \in \mathcal{N}_i(t)$  and  $\forall i \in \mathcal{V}$ . When  $\mathbf{x} \in \mathcal{Z}(\mathbf{w}) = \{\mathbf{x} \in \mathbb{R}^{nm} \mid \phi(\mathbf{x}, \mathbf{w}) \geq 0\} \subset \mathcal{C}$ , all robots have at least  $F'$  neighbors (by Lemma 2) and stay at least  $\Delta_d$  apart. Also let  $\partial \mathcal{Z}(\mathbf{w}) = \{\mathbf{x} \in \mathbb{R}^{nm} \mid \phi(\mathbf{x}, \mathbf{w}) = 0\}$ . Then, (11) is a CBF for system (2) if there exists an extended class  $\mathcal{K}_\infty$  function  $\alpha : \mathbb{R} \rightarrow \mathbb{R}$  such that

$$\sup_{\mathbf{u} \in \mathcal{U}} \left[ \sum_{i \in \mathcal{V}} \left( w_r \sum_{k \in \mathcal{B}_i(t)} E_k(\mathbf{x}) \frac{\partial h_k}{\partial x_i} + w_c \sum_{j \in \mathcal{N}_i(t)} E_{ij}^c(\mathbf{x}) \frac{\partial h_{ij}^{\text{col}}}{\partial x_i} \right) \dot{x}_i \right] \geq -\alpha(\phi(\mathbf{x}, \mathbf{w})), \quad (12)$$

$\forall \mathbf{x} \in \mathcal{Z}(\mathbf{w})$ . If such  $\alpha(\cdot)$  exists,  $\frac{\partial \phi}{\partial \mathbf{x}}(\mathbf{x}) \neq 0 \forall \mathbf{x} \in \partial \mathcal{Z}(\mathbf{w})$ , and  $\mathbf{x}(t_0) \in \mathcal{Z}(\mathbf{w})$ , then any Lipschitz continuous controller that satisfies (12) would make all robots maintain at least  $F'$

neighbors without collisions  $\forall t \geq t_0$ . The issue with (12) is that it is centralized; it depends on the states and dynamics of all agents. A robot  $i \in \mathcal{V}$  cannot compute its control input  $u_i \in \mathcal{U}_i$  while ensuring the satisfaction of (12) without considering the states and control inputs of all other robots. Various approaches exist to decouple CBF constraints into multiple distributed CBF constraints [2], [17], [26]. We adopt the strategy from [17], [26] by dividing the responsibility of satisfying (12) among robots.

### B. Constraint with Local Information

In this section, we decompose the CBF constraint (12) into  $n$  separate *distributed* constraints that only require local information. We first examine the terms  $\frac{\partial h_k}{\partial x_i}$ ,  $\frac{\partial h_{ij}^{\text{col}}}{\partial x_i}$ ,  $E_k(\mathbf{x})$ , and  $E_{ij}^c(\mathbf{x})$  in (12). Since robot  $i \in \mathcal{V}$  can measure physical states of its own and its neighbors, it can locally compute  $\frac{\partial h_k}{\partial x_i} \forall k \in \mathcal{B}_i(t)$  as well as  $\frac{\partial h_{ij}^{\text{col}}}{\partial x_i}$  and  $E_{ij}^c(\mathbf{x}) \forall j \in \mathcal{N}_i(t)$ . However, it may not have enough information to determine  $E_k(\mathbf{x})$  (since  $\mathcal{N}_k(t)$  and thus  $h_k(\mathbf{x})$  may be unknown) for  $k \in \mathcal{B}_i(t)$ . Thus, each robot constructs

$$\hat{h}_k(t) = c_k(t) - F', \quad (13)$$

which may or may not accurately represent  $h_k(\mathbf{x}(t))$  (9) at time  $t$ , depending on whether robot  $k$  updates  $c_k(t)$  according to (5). Let  $\hat{E}_k(t) = e^{-w_r \hat{h}_k(t)}$ . Also, for each  $i \in \mathcal{V}$ , let

$$\phi_i(\mathbf{x}, \mathbf{w}) = \frac{1}{n} - \frac{\sum_{k \in \mathcal{B}_i(t)} \hat{E}_k(t)}{F' + 1} - \frac{\sum_{j \in \mathcal{N}_i(t)} E_{ij}^c(\mathbf{x})}{2}, \quad (14)$$

which defines a set  $\mathcal{Z}_i(\mathbf{w}) = \{\mathbf{x} \in \mathbb{R}^{nm} \mid \phi_i(\mathbf{x}, \mathbf{w}) \geq 0\}$ . Due to the structure of (14),  $\phi_i(\mathbf{x}, \mathbf{w}) \geq 0$  at time  $t$  only when  $\hat{h}_k(t) > 0 \forall k \in \mathcal{B}_i(t)$  and  $h_{ij}^{\text{col}}(\mathbf{x}) > 0 \forall j \in \mathcal{N}_i(t)$  at  $t$ . Note  $\tilde{\mathcal{Z}}(\mathbf{w}) = \cap_{i \in \mathcal{V}} \mathcal{Z}_i(\mathbf{w})$  is not necessarily a subset of  $\mathcal{Z}(\mathbf{w})$  as  $\hat{h}_i(t)$  may not be equal to  $h_i(\mathbf{x}(t))$  for some  $i \in \mathcal{A}$ . Now, we present the distributed CBF constraint for agent  $i \in \mathcal{V}$ :

$$\underbrace{\left( w_r \sum_{k \in \mathcal{B}_i(t)} \hat{E}_k(t) \frac{\partial h_k}{\partial x_i} + w_c \sum_{j \in \mathcal{N}_i(t)} E_{ij}^c(\mathbf{x}) \frac{\partial h_{ij}^{\text{col}}}{\partial x_i} \right)}_{H_i(\mathbf{x})} \dot{x}_i \geq -\alpha(\phi_i(\mathbf{x}, \mathbf{w})), \quad (15)$$

where we assume  $H_i(\mathbf{x}) \neq 0, \forall \mathbf{x} \in \partial \tilde{\mathcal{Z}}_i(\mathbf{w}) = \{\mathbf{x} \in \tilde{\mathcal{Z}}(\mathbf{w}) \mid \phi_i(\mathbf{x}, \mathbf{w}) = 0\}, \forall i \in \mathcal{V}$ .

**Lemma 3.** Let  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$  be a network consisting of  $n$  agents with dynamics (1) and collective states  $\mathbf{x} \in \tilde{\mathcal{Z}}(\mathbf{w})$  for  $\mathbf{w} = [w_r, w_c] \in \mathbb{R}_+^2$  at time  $t \geq t_0$ . Let  $\alpha : \mathbb{R} \rightarrow \mathbb{R}$  be a linear extended class  $\mathcal{K}_\infty$  function. Suppose that in (11) and (12),  $E_i(\mathbf{x}(t))$  is replaced by  $\hat{E}_i(t)$ . Then, (15)  $\forall i \in \mathcal{V}$  is feasible. Also, if  $u_i(t)$  satisfies (15)  $\forall i \in \mathcal{V}$  at  $t$ ,  $\mathbf{u}(t)$  satisfies (12) at  $t$ .

*Proof.* First, we prove the feasibility of (15). Since  $\mathbf{x} \in \tilde{\mathcal{Z}}(\mathbf{w})$ , we know  $\phi_i(\mathbf{x}, \mathbf{w}) \geq 0 \forall i \in \mathcal{V}$ . Then,  $u_i(t) = 0$  satisfies (15)  $\forall i \in \mathcal{V}$ , regardless of  $\alpha(\cdot)$ . Now we prove that

satisfying (15)  $\forall i \in \mathcal{V}$  also implies the satisfaction of (12). Since  $|\mathcal{B}_k(t)| \geq F' + 1 \forall k \in \mathcal{B}_i(t), \forall i \in \mathcal{V}$ ,

$$\sum_{i \in \mathcal{V}} \hat{E}_i(t) \leq \sum_{i \in \mathcal{V}} \frac{\sum_{k \in \mathcal{B}_i(t)} \hat{E}_k(t)}{F' + 1}. \quad (16)$$

Also, because two agents  $i, j$  are involved in  $h_{ij}^{\text{col}}(\mathbf{x})$ ,

$$\sum_{(i,j) \in \mathcal{E}(t)} E_{ij}^c(\mathbf{x}) = \sum_{i \in \mathcal{V}} \frac{\sum_{j \in \mathcal{N}_i(t)} E_{ij}^c(\mathbf{x})}{2}. \quad (17)$$

Thus, combining (16) and (17), we get

$$\phi(\mathbf{x}, \mathbf{w}) \geq \sum_{i \in \mathcal{V}} \phi_i(\mathbf{x}, \mathbf{w}). \quad (18)$$

Therefore, with any linear function  $\alpha(\cdot)$ , we can conclude

$$\sup_{\mathbf{u} \in \mathcal{U}} \left[ \sum_{i \in \mathcal{V}} \left( w_r \sum_{k \in \mathcal{B}_i(t)} \hat{E}_k(t) \frac{\partial h_k}{\partial x_i} + w_c \sum_{j \in \mathcal{N}_i(t)} E_{ij}^c(\mathbf{x}) \frac{\partial h_{ij}^{\text{col}}}{\partial x_i} \right) \dot{x}_i \right] = \quad (19a)$$

$$\sum_{i \in \mathcal{V}} \sup_{u_i \in \mathcal{U}_i} \left[ \left( w_r \sum_{k \in \mathcal{B}_i(t)} \hat{E}_k(t) \frac{\partial h_k}{\partial x_i} + w_c \sum_{j \in \mathcal{N}_i(t)} E_{ij}^c(\mathbf{x}) \frac{\partial h_{ij}^{\text{col}}}{\partial x_i} \right) \dot{x}_i \right] \geq - \sum_{i \in \mathcal{V}} \alpha(\phi_i(\mathbf{x}, \mathbf{w})) \geq -\alpha(\phi(\mathbf{x}, \mathbf{w})), \quad (19b) \quad (19c)$$

which completes the proof.  $\square$

Lemma 3 shows that each robot  $i \in \mathcal{V}$  satisfying its own constraint (15) leads to all robots collectively satisfying (12) with  $h_k(\mathbf{x}(t))$  being replaced by  $\hat{h}_k(t)$ , regardless of whether  $\hat{h}_k(t) = h_k(\mathbf{x}(t))$  or not. We also highlight that all the components of (15) for robot  $i \in \mathcal{V}$  can be derived only using its local information. Using this fact, we now design our distributed controller.

### C. Controller Design

Now, we present our controller. Let  $u_{i,\text{des}}(x_i)$  be the desired controller of robot  $i$ . Our CBF-based Quadratic Program (CBF-QP) controller for robot  $i \in \mathcal{V}$  is:

$$u_i^*(\mathbf{x}) = \arg \min_{u_i \in \mathcal{U}_i} \|u_{i,\text{des}}(x_i) - u_i\|^2 \quad (20)$$

s.t. (15) is satisfied.

**Assumption 2.**  $u_i^*(\mathbf{x})$  (20) is locally Lipschitz continuous with respect to  $\mathbf{x}$ .

**Remark 2.** In general, CBF-QPs with input constraints do not yield locally Lipschitz continuous solutions, as the input constraints effectively act as additional constraints. The local Lipschitz continuity of QPs with multiple constraints have been studied in [27]–[29].

Now, we prove that our controller ensures both resilient consensus and physical safety under a specific class of malicious attacks on  $c_i(t)$ :

**Theorem 1.** Let all conditions in Lemma 2-3 and Assumption 1-2 hold. Suppose all misbehaving agents  $i \in \mathcal{A}$  update  $c_i(t)$  according to  $c_i(t) = \sum_{j \in \mathcal{N}_i(t)} a_{ij}(\mathbf{x}(t)) + \epsilon_i, \forall t \geq t_0$ , where  $\epsilon_i \in \mathbb{R}$ , such that  $\mathbf{x}(t_0) \in \tilde{\mathcal{Z}}(\mathbf{w})$ . If each robot  $i \in \mathcal{V}$  computes its control input  $u_i$  using (20), then  $\forall t \geq t_0$ , all robots move such that (a)  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \geq F'$ , (b) no collisions occur, and (c) resilient consensus is guaranteed.

*Proof. (Part 1)* By Assumption 1, all robots compute  $u_i(t)$  from (20)  $\forall t \geq t_0$ , which means  $\mathbf{u}(t)$  from (20)  $\forall i \in \mathcal{V}$  satisfies (12)  $\forall t \geq t_0$  by Lemma 3. Since (a)  $u_i^*(\mathbf{x}) \forall i \in \mathcal{V}$  are locally Lipschitz by Assumption 2 and feasible  $\forall \mathbf{x} \in \tilde{\mathcal{Z}}(\mathbf{w})$  by Lemma 3, (b)  $\mathbf{x}(t_0) \in \tilde{\mathcal{Z}}(\mathbf{w})$ , and (c)  $H_i(\mathbf{x}) \neq 0 \forall \mathbf{x} \in \partial \tilde{\mathcal{Z}}(\mathbf{w})$ , the controllers  $u_i^*(\mathbf{x}) \forall i \in \mathcal{V}$  together render  $\tilde{\mathcal{Z}}(\mathbf{w})$  forward invariant.

*(Part 2)* Because  $\tilde{\mathcal{Z}}(\mathbf{w})$  is rendered forward invariant,  $\forall t \geq t_0$  (i)  $h_i(\mathbf{x}(t)) = \hat{h}_i(t) > 0 \forall i \in \mathcal{L}$ , (ii)  $h_i(\mathbf{x}(t)) = \hat{h}_i(t) - \epsilon_i > 0 \forall i \in \mathcal{A}$ , and (iii)  $h_{ij}^{\text{col}}(\mathbf{x}(t)) > 0 \forall i \in \mathcal{V} \forall j \in \mathcal{N}_i(t)$ . Then, by Lemma 2, robots are guaranteed to move such that  $\delta_{\mathcal{L}}(\mathcal{G}(t)) \geq F'$  without collision  $\forall t \geq t_0$ . Also, by Proposition 1, resilient consensus is guaranteed.  $\square$

**Remark 3.** Our analysis in Theorem 1 assumes malicious robots conduct constant biased attacks on  $c_i(t)$  (but not necessarily on  $y_i(t)$ ), which are common forms of attacks due to their stealthiness [30]. However, these represent a specific class of malicious attacks. Extending our analysis to broader attack strategies is future work.

We also examine the performance of our controller (20) in the cases where all robots  $i \in \mathcal{V}$  correctly update connectivity levels  $c_i(t)$  according to the nominal update rule (5) below:

**Corollary 1.** Let all conditions in Theorem 1 hold. Suppose all robots  $i \in \mathcal{V}$  update  $c_i(t)$  according to (5), i.e.,  $\epsilon_i = 0, \forall i \in \mathcal{A}$ , such that  $\mathbf{x}(t_0) \in \tilde{\mathcal{Z}}(\mathbf{w})$ . If each robot  $i \in \mathcal{V}$  computes its control input  $u_i$  using (20), then  $\forall t \geq t_0$ , all robots move such that (a)  $\delta_{\min}(\mathcal{G}(t)) \geq F'$ , (b) no collisions occur, and (c) resilient consensus is guaranteed.

*Proof. (Part 1)* By the same argument as in (Part 1) of the proof of Theorem 1, the controllers  $u_i^*(\mathbf{x}) \forall i \in \mathcal{V}$  together render  $\tilde{\mathcal{Z}}(\mathbf{w})$  forward invariant.

*(Part 2)* Because  $\epsilon_i = 0 \forall i \in \mathcal{A}$  and  $\tilde{\mathcal{Z}}(\mathbf{w})$  is rendered forward invariant,  $\forall t \geq t_0$  (i)  $h_i(\mathbf{x}(t)) = \hat{h}_i(t) > 0 \forall i \in \mathcal{V}$  and (ii)  $h_{ij}^{\text{col}}(\mathbf{x}(t)) > 0 \forall i \in \mathcal{V} \forall j \in \mathcal{N}_i(t)$ . Then, by Lemma 2, robots are guaranteed to move such that  $\delta_{\min}(\mathcal{G}(t)) \geq F'$  without collision  $\forall t \geq t_0$ . Also, by Proposition 1, resilient consensus is guaranteed.  $\square$

**Remark 4.** As shown in Theorem 1 and Corollary 1, while resilient consensus and collision avoidance are always ensured, the performance of our designed controller (20) (i.e., the overall connectivity of the network) depends on the behavior of malicious robots. If all robots  $i \in \mathcal{V}$  update their connectivity values  $c_i(t)$  according to (5), then all robots are guaranteed to maintain at least  $F'$  neighbors while avoiding collisions. However, if malicious robots  $i \in \mathcal{A}$  manipulate  $c_i(t)$ , they may fail to maintain  $F'$  neighbors.



This also affects how aggressively/conservatively robots can move with our controller. We examine this effect more closely in our simulation results.

## V. SIMULATION RESULTS

Here, we show the effectiveness of our controller (20) through simulations in a 2D plane. We consider a time-varying network  $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ , consisting of  $n = 11$  robots, where edges are formed based on robots' relative distances with a communication range  $R = 3$  m. Each robot has dynamics (1) and an input set  $\mathcal{U}_i = [-1.5, 15] \times [-1.5, 1.5]$ . They are simulated as point masses and to maintain a minimum distance of  $\Delta_d = 0.3$  m. Each normal robot  $i \in \mathcal{L}$  updates its connectivity value  $c_i(t)$  according to (4) and shares it with its neighbors every  $\tau_1 = 0.005$  seconds. It also starts with a random consensus state  $y_i(t) \in [-500, 500]$  at  $t = t_0$ , which it shares and updates every  $\tau_2 = 0.1$  seconds using W-MSR algorithm thereafter. We examine three scenarios, each involving  $F = 2$  malicious robots  $i \in \mathcal{A}$  following different update rules for  $c_i(t)$ :

**1. Nominal Updating:** Follow the nominal law (4) to update:

$$c_i(t) = \sum_{j \in \mathcal{N}_i(t_l)} a_{ij}(\mathbf{x}(t_l)), \quad \forall t \in [t_l, t_{l+1}), \quad \forall l \in \mathbb{Z}_{\geq 0}.$$

**2. Understating:** Understate their connectivity levels:

$$c_i(t) = \sum_{j \in \mathcal{N}_i(t_l)} a_{ij}(\mathbf{x}(t_l)) - 2.5, \quad \forall t \in [t_l, t_{l+1}), \quad \forall l \in \mathbb{Z}_{\geq 0}.$$

**3. Overstating:** Overstate their connectivity levels:

$$c_i(t) = \sum_{j \in \mathcal{N}_i(t_l)} a_{ij}(\mathbf{x}(t_l)) + 3.5, \quad \forall t \in [t_l, t_{l+1}), \quad \forall l \in \mathbb{Z}_{\geq 0}.$$

In all cases, all malicious robots randomly choose and share  $y_i(t) \in [-500, 500]$  every  $\tau_2$  seconds. By Proposition 1, all normal robots have to maintain  $F' = 2 + \lfloor \frac{11}{2} \rfloor = 7$  or more neighbors to ensure resilient consensus. We evaluate our controller by observing whether robots can achieve resilient consensus even when the robots' desired control inputs directly conflict with maintaining network connectivity, despite malicious robots' adversarial behaviors. With the desired controller  $u_{i,\text{des}}(x_i) = \frac{v(x_i)}{\|v(x_i)\|}$  for each robot  $i \in \mathcal{V}$ , where

$$v(x_i) = \begin{cases} (-1)^i \cdot [100, 0]^T - x_i & \text{if } i \in \{1, \dots, 5\}, \\ (-1)^i \cdot [0, 100]^T - x_i & \text{if } i \in \{6, \dots, 11\}, \end{cases}$$

robots will move in four different directions, which would break the network and prevent resilient consensus. Each simulation begins with every robot having at least seven neighbors, as shown in Fig. 1 (a), where normal and malicious robots are blue and red, respectively. Links between normal robots are depicted as gray dotted lines, while links involving malicious robots are shown as red dotted lines.

The robot formations at  $t = 10$  seconds for each scenario are shown in Fig. 1 (b)-(d). Regardless of whether malicious robots  $i \in \mathcal{A}$  update  $c_i(t)$  according to (4), robots eventually stop dispersing to allow normal robots maintain at least seven neighbors while avoiding collisions, validating Theorem 1

and Corollary 1. Fig. 2 (a)-(c) show evolution of the degree constraint functions (9) for each scenario. The results also highlight the effect of malicious behaviors on robots' movement, which is discussed in Remark 4. Comparing Fig. 1 (b) and (c), we observe that understating behaviors impose more connectivity constraints for normal robots to remain close to malicious agents, leading to a more compact network. Conversely, as seen in Fig. 1 (b) and (d), overstating behaviors relax these constraints, allowing normal agents to move more freely and form a more dispersed network. Notably, one malicious robot fails to maintain seven neighbors (shown in Fig. 2 (c)). Importantly, these variations in updates do not compromise resilient consensus (as shown in Fig. 2 (d)-(f)), demonstrating the effectiveness of our controller.

## VI. CONCLUSION

In this work, we propose a distributed CBF-based controller that guarantees resilient consensus among robots during navigation without relying on fixed network topologies. We first establish a sufficient condition for resilient consensus based solely on the minimum degree of normal agents in a network. Leveraging this condition, we design a controller that enables robots to make movement decisions using only local information, ensuring resilient consensus in a time-varying network. Furthermore, our current analysis assumes continuous connectivity information and a specific form of malicious attack, while real systems operate with discrete updates and potentially varied threats. We aim to address these issues in future extensions.

## REFERENCES

- [1] H. Zhu, J. Juhl, L. Ferranti, and J. Alonso-Mora, "Distributed multi-robot formation splitting and merging in dynamic environments," in *International Conference on Robotics and Automation (ICRA)*, 2019, pp. 9080–9086.
- [2] X. Tan and D. V. Dimarogonas, "Distributed implementation of control barrier functions for multi-agent systems," *IEEE Control Systems Letters*, vol. 6, pp. 1879–1884, 2022.
- [3] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 4, pp. 766–781, 2013.
- [4] Y. Wang, H. Ishii, F. Bonnet, and X. Défago, "Resilient real-valued consensus in spite of mobile malicious agents on directed graphs," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 3, pp. 586–603, 2022.
- [5] H. Lee and D. Panagou, "Construction of the sparsest maximally  $r$ -robust graphs," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 170–177.
- [6] —, "Minimal construction of graphs with maximum robustness," *arXiv preprint arXiv:2507.00415*, 2025.
- [7] H. J. LeBlanc and X. D. Koutsoukos, "Low complexity resilient consensus in networked multi-agent systems with adversaries," in *Proceedings of the 15th ACM International Conference on Hybrid Systems: Computation and Control*. New York, NY, USA: Association for Computing Machinery, 2012, p. 5–14.
- [8] H. Zhang, E. Fata, and S. Sundaram, "A notion of robustness in complex networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 310–320, 2015.
- [9] D. Saldaña, A. Prorok, S. Sundaram, M. F. M. Campos, and V. Kumar, "Resilient consensus for time-varying networks of dynamic agents," in *2017 American Control Conference (ACC)*, 2017, pp. 252–258.
- [10] J. Usevitch and D. Panagou, "Resilient leader-follower consensus to arbitrary reference values in time-varying graphs," *IEEE Transactions on Automatic Control*, vol. 65, pp. 1755–1762, 2020.

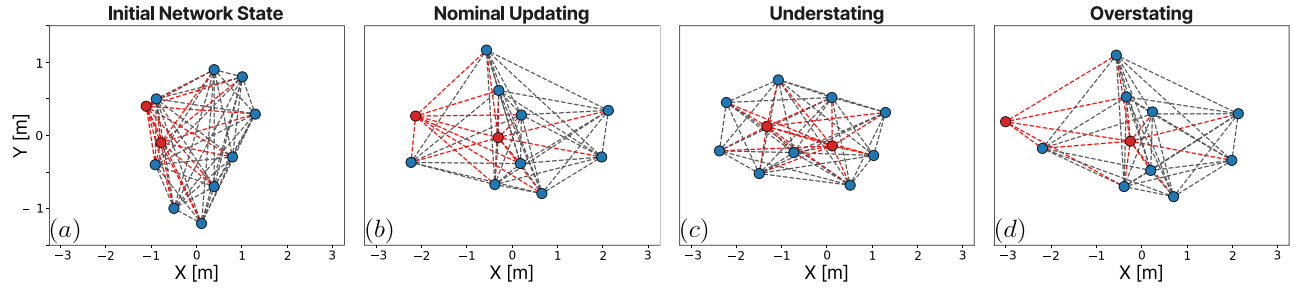


Fig. 1. (a) shows the initial robot formation at  $t_0 = 0$ , while (b)–(d) depict formations at  $t = 10$  seconds for nominal updating, understating, and overstating scenarios, respectively. Gray dotted lines are connections between normal robots, while red dotted lines are those involving malicious robots.

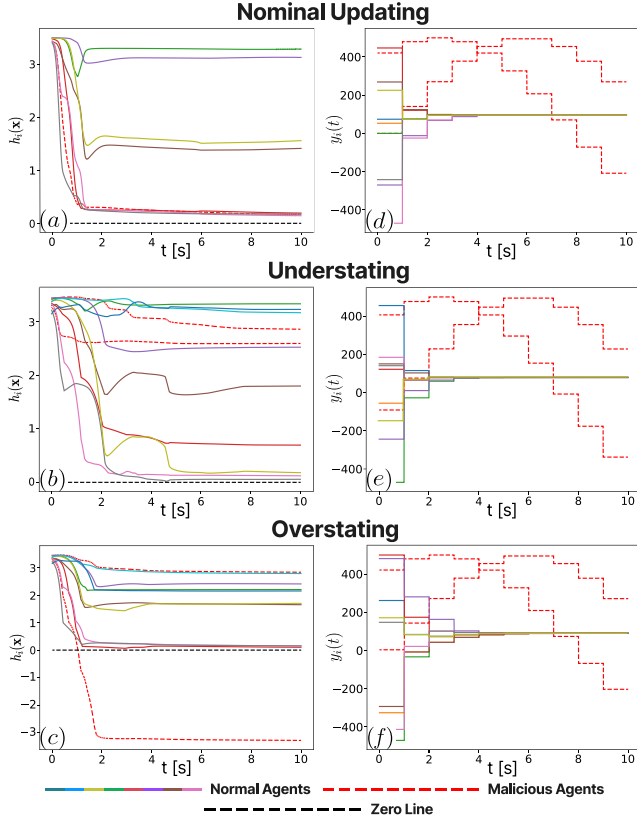


Fig. 2. (a)–(c) show the robots’ actual values of (9), while (d)–(f) present the robots’ consensus states, both for the simulations of nominal updating, understating, and overstating scenarios for  $t \in [0, 10]$ , respectively.

[11] L. Guerrero-Bonilla and V. Kumar, “Realization of  $r$ -robust formations in the plane using control barrier functions,” *IEEE Control Systems Letters*, vol. 4, no. 2, pp. 343–348, 2020.

[12] K. Saulnier, D. Saldaña, A. Prorok, G. J. Pappas, and V. Kumar, “Resilient flocking for mobile robot teams,” *IEEE Robotics and Automation Letters*, vol. 2, pp. 1039–1046, 2017.

[13] M. Cavorsi, B. Capelli, and S. Gil, “Multi-robot adversarial resilience using control barrier functions,” *Robotics: Science and Systems*, 2022.

[14] Z. Zhang, Y. Wu, J. Jiang, N. Zheng, and W. Meng, “Realization of robust formation for multi-uav systems using control barrier functions,” *Unmanned Systems*, vol. 0, no. 0, pp. 1–15, 2024.

[15] H. Lee and D. Panagou, “Maintaining strong  $r$ -robustness in reconfigurable multi-robot networks using control barrier functions,” in *IEEE International Conference on Robotics and Automation*, 2025.

[16] A. Ames, S. D. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, “Control barrier functions: Theory and applications,” *2019 18th European Control Conference (ECC)*, pp. 3420–3431, 2019.

[17] P. Bhatia, S. B. Roy, P. Sujit, L. M. Alvarez, and A. McFadyen,

“Decentralized Connectivity Maintenance for Multi-Agent Systems Using Control Barrier Functions,” in *2024 International Conference on Unmanned Aircraft Systems (ICUAS)*. Chania - Crete, Greece: IEEE, Jun. 2024, pp. 955–962.

[18] N. De Carli, P. Salaris, and P. R. Giordano, “Distributed Control Barrier Functions for Global Connectivity Maintenance,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. Yokohama, Japan: IEEE, May 2024, pp. 12 048–12 054.

[19] B. Capelli and L. Sabattini, “Connectivity maintenance: Global and optimized approach through control barrier functions,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 5590–5596.

[20] W. Ren and R. Beard, “Consensus seeking in multiagent systems under dynamically changing interaction topologies,” *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[21] H. J. LeBlanc, *Resilient cooperative control of networked multi-agent systems*. Vanderbilt University, 2012.

[22] M. Fiedler, “Algebraic connectivity of graphs,” *Czechoslovak Mathematical Journal*, vol. 23, no. 2, pp. 298–305, 1973.

[23] O. Dagdeviren and V. Khalilpour Akram, “KEIP: a distributed  $k$ -connectivity estimation algorithm based on independent paths for wireless sensor networks,” *Wireless Networks*, vol. 25, no. 8, pp. 4479–4491, Nov. 2019.

[24] M. Black and D. Panagou, “Adaptation for validation of consolidated control barrier functions,” in *IEEE Conference on Decision and Control (CDC)*, 2023, pp. 751–757.

[25] P. Glotfelter, J. Cortés, and M. Egerstedt, “Nonsmooth barrier functions with applications to multi-robot systems,” *IEEE Control Systems Letters*, vol. 1, no. 2, pp. 310–315, 2017.

[26] M. Jankovic, M. Santillo, and Y. Wang, “Multiagent Systems With CBF-Based Controllers: Collision Avoidance and Liveness From Instability,” *IEEE Transactions on Control Systems Technology*, vol. 32, pp. 705–712, Mar. 2024.

[27] G. Still, “Lectures on parametric optimization: An introduction,” *Optimization Online*, p. 2, 2018.

[28] P. Mestres, A. Allibhoy, and J. Cortés, “Regularity properties of optimization-based controllers,” *European Journal of Control*, p. 101098, Jan. 2025.

[29] D. R. Agrawal, H. Lee, and D. Panagou, “Reformulations of quadratic programs for lipschitz continuity,” *arXiv preprint arXiv:2508.18530*, 2025.

[30] A. Teixeira, G. Dán, H. Sandberg, and K. H. Johansson, “A cyber security study of a scada energy management system: Stealthy deception attacks on the state estimator,” *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 11 271–11 277, 2011, 18th IFAC World Congress.