Thesis Summary : Optimal Scaled Attacks on Consensus-based Formation Control

Sreejeet Maity[1] and Vaibhav Katewa[2]

Abstract— **Formation control is one of the primal control methodologies in the broad domain of multi-agent systems(MASs). It is extensively used in multi-agent robotics (MARs), unmanned autonomous vehicles, and satellite networks. Recently, to minimize human intervention, global superpowers have been heavily investing in the defense sector for deploying uncrewed aerial/ground vehicles to carry out surveillance in potentially challenging environments. Apart from physical robustness, one primary challenge is making the deployed MAS resilient against cyber-attacks. Despite its widespread relevance, minimal literature is available on the security aspects of formation control. Our research primarily aims to develop and analyze potential security aspects of formation control on a Multi-Robot System(MRS). In this paper, we propose a novel attack capable of influencing one or more agents of an MRS to distort the desired formation in a finite amount of time(distortion time). In the presence of our proposed attack, we investigated the distortion time through the spectral properties of the underlying network topology(Agent Connectivity). We also looked for a possible trade-off between the underlying network topology and the extent of distortion(Formation Error). Lastly, the experimental results in the presence/absence of our proposed attack illustrate the attack severity in terms of formation error and distortion time.**

I. INTRODUCTION

In recent years, interest in developing various multi-agent systems has gained massive momentum [3],[4]. MAS consists of one or more cooperative intelligent agents, locally communicating amongst themselves to achieve an allotted task [1]. MRS is also noted to be significantly reliable, efficient, and more robust than a single robot.

Most of the problem in the field of MRS requires the agents to maintain some formation with respect to each other. As a result, cutting-edge research on theoretical and applied aspects of formation control seems ever-growing [1]. From industrial automation to satellite networks, connected autonomous vehicles(CAVs) to unmanned aerial/underwater vehicles, formation control plays a cardinal role. The agents in an MRS are expected to maintain a desired distance from each other to avoid collisions and major mishaps [4], [6]. As a result, in MRS, agents must communicate their spatial or any other relevant parameters with the other agents in the group. The underlying topology that governs this information exchange among the agents plays a requisite role in determining the extent of cooperation among the agents.

As far as applications are concerned, formation control is widely used on unmanned aerial/ underwater vehicles to carry out surveillance, searching, and other things in a systematic manner. Similarly, in industrial applications such as payload transfer or object transportation, maintaining a formation is crucial to properly balance the load [5]. In some instances, the agents might also require strategic switching formation to cope with the challenges faced during the transit. In a nutshell, the primal objective of formation control is to direct the agents to maintain a desired geometric structure and enforce them to traverse as a unit [4],[5].

Going by the literature [6], formation control of an MRS is ensured by regular intra-agent information exchange. Since mobility is an important facet of MRSs, wired communications are not preferred despite potential security advantages. Since broadcast nature is inherent to wireless communications, an intruder can easily access and manipulate the exchanged pieces of information [2].

MRSs relying on wireless communication are strongly susceptible to cyber attacks. The current literature is built chiefly upon strategic attack schemes on the communication channels. For example, Shames et al. [7] dealt with the sequential detection and gradual omission of a malicious agent from the entire group. Resilient and robust consensus protocols were studied by Koutsoukus and Le Blanc [8]. Feng et al. [9], [25] achieved secure consensus schemes and overcame adversaries by using switching topology in the deterministic setup and in a non-deterministic arrangement (governed by the Markov process), respectively. An intruder can also affect the system to repeat control signals (replay attacks) maliciously in a finite time step. The receding horizon control method was adopted by Zhu [10] to address such replay attacks. Recently, Al Yassin [17] and F. Louatti [18] adopted learning-based strategies to detect and restore a comprehensive class of replay attacks in multi-agent setups. False-data injection models [11], [13] are another relevant cyber attack having the potential to manipulate and distort trajectories and final states of individual agents without destabilizing the system. If the attack is appropriately designed, real-time detection of such a cyber attack might be difficult. Zhang et al. [14] achieved significant resiliency in a networked control system by adopting Denial-of-Service (DoS) attack. However, the major breakthrough of [12] lies in utilizing a functional loop method to derive necessary and sufficient conditions for closed-loop stability under the influence of a DoS attack.Shao and Ye [22] designed secure

*This work was not supported by any organization

[1]Vaibhav Katewa is with the Robert Bosch Centre for Cyber-Physical Systems, Indian Institute of Science, Bengaluru vkatewa@iisc.ac.in

[2]Sreejeet Maity is with the Robert Bosch Centre for Cyber-Physical Systems, Indian Institute of Science, Bengaluru sreejeetm@iisc.ac.in

control mechanisms to nullify the adverse effects in the presence of DoS. Apart from that, [22],[26] tried to alleviate related attacks and achieved resiliency in event-triggered communication frameworks. It is increasingly evident that researchers are getting more interested in MRS (or MAS in general) security aspects. However, most of the present literature is biased toward attack detection and nullification of the resultant adversaries . To date, researchers have yet to consider attack design and analysis thoroughly.

II. Consensus Based Formation Control for MAS

In the communication framework, we assume the agents can convey their positional information with their respective neighbors concerning a reference coordinate system. We proposed a one-shot displacement-based consensus protocol that enables a group of N agents to achieve a user-defined formation. $x_i(t) \in \mathbb{R}^d$ conceals the state of the i^{th} agent in time t. This paper will refer to $x_i(t)$ as x_i. Similarly, $x_i^{(1)} \in \mathbb{R}$ denotes the first component of $x_i \in \mathbb{R}^d$. The agent interactions represent the underlying network topology and can be modeled as a directed or undirected graph $G = \{V, \mathcal{E}\}$. Where $V \in \{1, n\}$ is the set of all vertices/agents and $\mathcal{E}$ is the edges that connect V. Lastly, N_i is the number of neighbors of i^{th} agent. The standard consensus-based protocol for the j^{th} ($j \in \{1, d\}$) component of x_i is defined as follows :

$$\dot{x_i}^{(j)} = \sum_{j \in N_i} a_{ij}(x_j^{(j)} - x_i^{(j)}) \tag{1}$$

Here, $a_{ij} \geq 0$ represents the weight of the edge between the i^{th} and j^{th} agent. In the non-weighted case, a_{ij} equals 1. And, a_{ij} equals to a_{ji} in the un-directed case.

We can modify the standard protocol given in (1) to incorporate consensus-based formation control (2). Assuming, δ_{ij} (equals to δ_i - δ_j) to be the desired inter-agent displacement between the i^{th} agent and the j^{th} agent, where δ_i is the displacement vector directed towards the i^{th} agent from the origin. u_i is the control input for the i^{th} agent. We assume the underlying topology of (2) is weighted and directed (a_{ij} might not be equal to a_{ji}). Taking geometrical constraints into account, we need to incorporate the continuity condition, $\delta_{ij}^{(j)} + \delta_{jk}^{(j)} = \delta_{ik}^{(j)}\ \forall i, j, k \in \{1, n\}$.

$$u_i^{(j)} = \dot{x_i}^{(j)} = \sum_{j \in N_i} a_{ij}(x_j^{(j)} - x_i^{(j)} - \delta_{ji}^{(j)}) \tag{2}$$

A MAS consisting of N agents will achieve formation control under the influence of Equation (2) given the underlying topology G contains one or more Globally Reachable Node(GRN)[26]. Figure 2 represents two desired formations using 8 and 5 agents, respectively. A more compact way to write equation (2) is to introduce the Laplacian matrix ($L \in \mathbb{R}^n$) such that

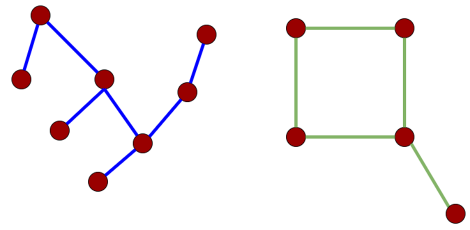

Fig. 1. Formation Control in MAS with 8 and 5 agents respectively

$$\dot{x} = -(L \otimes I_d)x + \bar{\delta} \tag{3}$$

Here, $x = [x_1^T\ x_2^T\ ...\ x_n^T]^T \in \mathbb{R}^{nd}$ and $\bar{\delta} \in \mathbb{R}^{nd}$ contains the conjugated information of the formation constraints. We can say the $[i, i+d-1]^{th}$ entry of $\bar{\delta}$ stands for $-\sum_{j \in N_i} \delta_{ji}$, and $\delta_{ji} \in \mathbb{R}^d$. Now, since the axis is decoupled, we can solve the consensus problem for each component and stack them in a pertinent manner to get the desired result.

A. Criteria for Convergence

The following convergence criteria for all $i, j \in \{1, n\}$ is initially defined for our problem of interest:

$$\lim_{t \to \infty} x_i(t) - x_j(t) = \delta_{ij} \tag{4}$$

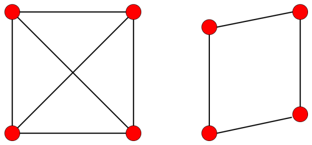

Fig. 2. Criteria influencing Multi-agent Formation Control with 4 agents

B. Formation control in the non-compromised scenarios

Lemma 1: Considering δ_{ij} as the steady displacement vector of the i^{th} agent from the j^{th} agent, there exists a unique vector w satisfying $\mathbb{1}^T w = 1$ such that solving equation (2) gives us the following result **R1**:

R1) $x_i(\infty) - x_j(\infty) = \delta_{ij}$

Proof : Let $z_i^{(j)} = x_i^{(j)}$ - $\delta_i^{(j)}$. We can rewrite $\dot{x_i}^{(j)} = \sum_{j \in N_i} a_{ij}(x_j^{(j)} - x_i^{(j)} - \delta_{ji}^{(j)})$ as $\dot{x_i}^{(j)} = \sum_{j \in N_i} a_{ij}(x_j^{(j)} - \delta_j^{(j)} - x_i^{(j)} + \delta_i^{(j)})$ or $\dot{x_i}^j = \sum_{j \in N_i} a_{ij}(z_j^{(j)} - z_i^{(j)})$. Since, $\dot{z_i}^{(j)} = \dot{x_i}^{(j)}$ we can write $\dot{z_i}^{(j)}$ =$\sum_{j \in N_i} a_{ij}(z_j^{(j)} - z_i^{(j)})$.Now, we can write $z_i(\infty)^{(j)}$=$w^T z(0)^{(j)}$, where w^T is the left dominant eigenvector (left eigenvector related to the highest eigenvalue) of the $-L$ and $z(0)^{(j)}$ are the initial values of $z^{(j)}$. Now, since $z_i^{(j)} = x_i^{(j)} - \delta_i^{(j)}$, we can write $x_i^{(j)} = w^T z(0)^{(j)} + \delta_i^{(j)}$. We can write $\delta^{(j)} = [\delta_1^{(j)}\ \delta_2^{(j)}\ ...\ \delta_n^{(j)}]^T$ In the next step, we rewrite the following expression by writing $x_i^{(j)} = w^T x(0)^{(j)} + \delta_i^{(j)} - w^T \delta^{(j)}$. Since, $\mathbb{1}^T w = 1$, we can take write components such as $x_i^{(j)} = w^T x(0)^{(j)} + w^T(\mathbb{1}_n \delta_i^{(j)} - \delta^{(j)})$. Subsequently, we can write the steady-state position of the i^{th} agent as :

$$x_i(\infty) = w^T[\bullet]X(0) + w^T[\bullet]\delta^{X_i(0)}$$

$$x_i(\infty) = w^T[\bullet](X(0) + \delta^{X_i(0)})$$

Where, $X(0) = [x(0)^{1^T} \; x(0)^{2^T} \; ... \; x(0)^{d^T}]^T$, $\delta^{X_i(0)} = [(\mathbb{1}_n\delta_i^{(1)} - \delta^{(1)})^T \; (\mathbb{1}_n\delta_i^{(2)} - \delta^{(2)})^T \; ... \; (\mathbb{1}_n\delta_i^{(d)} - \delta^{(d)})^T]^T$ and, [•] stands for penetrating-dot product. Similarly, we can proceed to write the j^{th} agent's final position as :

$$x_j(\infty) = w^T[\bullet](X(0) + \delta^{X_i(0)})$$

Given $w^T\mathbb{1}_n$=1, the steady state difference between the i^{th} and j^{th} agent is as follows :

$$x_i(\infty) - x_j(\infty) = w^T[\bullet](\delta^{X_i(0)} - \delta^{X_j(0)})$$

$$x_i(\infty) - x_j(\infty) = \delta_{ij}$$

Hence, **R1** is proved.

Corollary 1: If the underlying topology (G) is weighted and undirected, (2) solves the consensus problem when (G) is connected and gives us **P1** and **P2**:

P1) $x_i(\infty) = \frac{\mathbb{1}_n}{n}[\bullet](\delta^{X_i(0)} - \delta^{X_j(0)})$.

P2) $x_i(\infty) - x_j(\infty) = \delta_i - \delta_j = \delta_{ij}$

Proof : Substituting the undirected edges($\{i, j\}$) of G by two ordered pairs $\{i, j\}$ and $\{j, i\}$ respectively, we get an equivalent directed graph (G_1). This graph (G_1) is weight balanced. Hence, proceeding similarly as **Lemma 1** and using the results of [27], **P1** and **P2** can be proved.

III. Proposed attack on a given MRS

Attack design and analysis are critical while analyzing MRS. However, the literature has been biased toward attack detection and prevention [23]. Attacks on MRS can be classified into two broad categories: A) **Attacks on communication channels** and B) **Attacks on agents/nodes**. The attacks are designed to alter or manipulate agent trajectories, final positions, speed, other positional parameters, etc. We fundamentally assume the attacker can access communication channels and agents and store relevant information. Our proposed attack manipulates an agent/subset of MRS agents to share false data among the non-compromised peers. On top of that, the compromised agent also gets updated according to the manipulated data. In formation control problems, one obvious choice for false data injection is to make the compromised agent share the wrong positional information. [23] reviewed both time-varying and non-varying displacement-bias attacks and jamming attacks. However, displacement-bias attacks might prevent the agents from achieving the desired formation but cannot directly influence the distortion time. To address the shortcomings of [27], we propose an alternative attack where the compromised agent/agents share scaled positional information with the rest of the un-compromised team members. On top of that, the compromised agents are also updating themselves with the same information. Throughout this paper, we will adopt the term **SCALED ATTACK** for this. It is however essential to write the modified governing equation under the influence of the attack.

A. Modified Laplacian in the presence of scaled attack

In the previously proposed consensus protocol, the Laplacian will be modified under the influence of the scaled attack. Since it shares scaled positional information with the un-compromised team members and updates itself with the same amount, scaling i^{th} column by a constant $\alpha > 1$ should imply the i^{th} agent is compromised. Thus, if we define a diagonal matrix $D \in \mathbb{R}^n$ with i^{th} entry α_i , the modified system matrix is as follows :

$$\dot{x} = -(LD \otimes I_d)x + \bar{\delta} \tag{5}$$

When more than one agent is compromised, scaled attacks can be suitably redefined by increasing the number of non-unit entries in the diagonal matrix. The structure of D is similar to that of D given below, α_i equals to 1 when the i^{th} agent is not compromised or has no outgoing edge. Similarly, we take $\alpha_i \in \mathbb{R}^+$ when it is attacked and is sharing state information by scaling it with α_i. In a directed graph, the attack will only affect the system when the attacking node/nodes have outgoing edge/edges. In that case, α_i should be equal to 1.

$$D = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix}$$

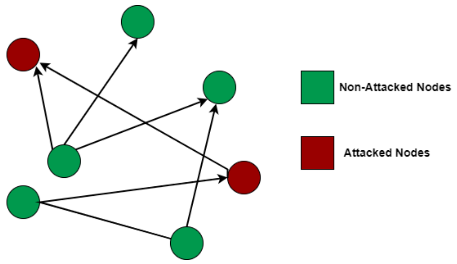

Fig. 3. MAS with two attacked agent (coloured red) and 5 non-attacked agent(coloured green) trying to achieve a formation.

B. Convergence guarantee under the scaled attack protocol

Theorem 1: The attack protocol defined in (4) gives the following two results, namely **R1** and **R2** $\Longleftrightarrow$ **R3**:

R1) The eigenvalues $\mu_i(LD) \;\; > 0 \; \forall \; i \in \{2, n\}$ of the compromised system, and $\mu_1(LD)$ = 0.

R2) The underlying topology stays invariant under the proposed attack.

R3) Alternatively, $\mathbb{1}_{sgn\{[L\otimes I_d]_{ij}\}=sgn\{[LD\otimes I_d]_{ij}\}\geq 0} = 1$, where, $\mathbb{1}$ and sgn denote the standard indicator and sign function respectively.

Proof : For proving **R1**, we will start with the structure of DL. The structures of L and DL are as follows

$$L = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix}, DL = \begin{pmatrix} \alpha_1 \cdot R_1 \\ \alpha_2 \cdot R_2 \\ \vdots \\ \alpha_n \cdot R_n \end{pmatrix}$$

Here, R_i represents the i^{th} row of L. Since, $\sum_j [R_i]_{ij} = 0$, we can also say $\sum_j \alpha_i [R_i]_{ij} = \alpha_i \sum_j [R_i]_{ij} = 0$. Hence, $D \cdot L$ is a Laplacian Matrix. We know $\mu_i(DL) \geq 0$ $\forall i \in \{1, n\}$. Since D is invertible , we can see $LD = D^{-1}(DL)D$ by similarity transformation. Hence, we can say $\mu_i(DL)$=$\mu_i(LD)$ $\forall$ $i \in \{1, n\}$. Since, DL is also a Laplacian by [26], and is connected, the 0 eigenvalue of DL is simple and rest $\mu_i(DL)$ are > 0. Hence, similarly transformed $\mu_i(LD)$ also possess the same. Hence, **R1** is proved. Going by our definition, $\alpha_i > 0$, hence $sgn(L) = sgn(LD)$ for any $\alpha \in \mathbb{R}^+$ proving **R3**. Continuing from **R3**, we can see the weights of the edges have changed, but none of the weights are going to 0 or changing sign. Hence the underlying topology should remain invariant, proving **R2** and vice-versa.

IV. Formation Control under the influence of the scaled attack

We are modifying the proposed protocol to incorporate the scaled attack dynamics to analyze the extent of distortion. Expanding Eq. 4, we get :

$$\dot{x_i}^{(j)} = \sum_{j \in N_i} a_{ij}(\alpha_j x_j^{(j)} - \alpha_i x_i^{(j)} - \delta_{ji}^{(j)}) \tag{6}$$

A. Final position of the agents under scaled attacks

Lemma 2: Considering δ_{ij} as the steady state displacement vector of the i^{th} agent from the j^{th} agent, there exists a unique vector $\bar{w}$ satisfying $\mathbb{1}^T \bar{w} = 1$ such that solving equation (6) gives us the following result **R1**:
R1) $\alpha_i \cdot x_i(\infty) - \alpha_j \cdot x_j(\infty)$ = δ_{ij}

Proof: We start our analysis by considering $z_i^{(j)} = \alpha_i x_i^{(j)} - \delta_i^{(j)}$. Thus, $\dot{z_i}^{(j)}$ equals to $\alpha_i \dot{x_i}^{(j)}$. As a result of that, Eq. 6 gets modified as $\frac{\dot{z_i}^{(j)}}{\alpha_i}$ = $\sum_{j \in N_i} a_{ij}(z_j^{(j)} - z_i^{(j)})$ or equivalently,

$$\dot{z_i}^{(j)} = \sum_{j \in N_i} \alpha_i a_{ij}(z_j^{(j)} - z_i^{(j)}) \tag{7}$$

Similarly we could see that $z_i^{(j)}(\infty)$=$\bar{w}^T z(0)^{(j)}$, where $\bar{w}^T = \beta \cdot \tilde{w}^T$. Taking $\alpha = [\alpha_1 \;\; \alpha_2 \;\; ... \;\; \alpha_n]^T$, and $\beta = \frac{1}{\sum_i \frac{w_i}{\alpha_i}}$ we can write:

$$x_i^{(j)} = \beta \left[\frac{\tilde{w}^T}{\alpha_i}(\alpha \odot x(0)^{(j)}) + \frac{\tilde{w}^T}{\alpha_i} K_i^j \right]$$

Where, $K_i^j = \left(\mathbb{1}_n \delta_i^{(j)} - \delta^{(j)} \right)$.Proceeding by the lines of Lemma 1:

$$x_i = \beta \left[\frac{\tilde{w}^T}{\alpha_i} [\bullet] \{ (\alpha \otimes \mathbb{1}_d) \odot X(0) + \delta^{X_i(0)} \} \right]$$

Similarly,

$$(\alpha_i \cdot x_i - \alpha_j \cdot x_j) = \beta \tilde{w}^T \mathbb{1}_n \delta_{ij} = \delta_{ij}$$

Hence, Lemma 2 is proved.
From now on, we will use $\tilde{x}_i$ to represent steady state of the i^{th} agent under the attacked scenario to differentiate between the non- compromised and compromised case.
For the expression of the eigenvector corresponding to 0 eigenvalues in the attacked case, we provide the following Lemma 3

Lemma 3: In the compromised case (6), the left eigenvector of the system (LD) corresponding to the zero eigenvalues of algebraic multiplicity(A.M) $= 1$ is $\bar{w} = \beta \tilde{w}$. Here $\beta = \frac{1}{\sum_i \frac{w_i}{\alpha_i}}$ and $\tilde{w} = \frac{1}{\alpha} \otimes w$, where w is the eigenvector satisfying the same property in the non-compromised case and $\frac{1}{\alpha} = [\frac{1}{\alpha_1} \;\; \frac{1}{\alpha_2} \cdots \frac{1}{\alpha_n}]^T$.

Proof: From (7), we see the row-wise entries of the uncompromised Laplacian matrix (L) get scaled by α_i, where 'α_i' corresponds to the scaling of the i^{th} row. From Theorem 1, we know the attack does not change the underlying topology. Hence, the unique eigenvector corresponding to 0 eigenvalues in the attacked scenario is $\tilde{w} = \left[\frac{w_1}{\alpha_1} \frac{w_2}{\alpha_2} \cdots \frac{w_n}{\alpha_n} \right]^T$, since $\tilde{w}^T(DL) = 0$. Since, we also have to ensure $\tilde{w}^T \mathbb{1}_n$=1, we introduced the scaling factor β such that $\beta \tilde{w}^T(DL)$ = 0.

V. Attack analysis: Distortion Time

As evident from (7), the proposed attack (5) alters the dynamics of the compromised system. In the attacked scenario, the time taken for distortion is one of the two principal pillars for analyzing attack severity. We are aware from [28],[27] that the time taken for convergence is solely dependent on the second smallest eigenvalue of the system matrix. Let $\mu_i(\geq 0)$ and $\lambda_i(\geq 0)$ $\forall i \in \{1, n\}$ be the eigenvalues of the compromised system and non-compromised system, respectively. Index $i = 2$ and $i = n$ refer to the second-smallest and largest eigenvalue in both cases.

A. Type 1 and Type 2 Convergence

Keeping the problem statement in mind, in this paper we introduce two convergence criterias namely Type-1 Convergence and Type-2 Convergence. We will state them as follows :
Definition 1: Assume there exist two systems A and B which satisfies the generic differential equations $\dot{x} = -Ax + \delta$ and $\dot{x} = -Bx + \delta$ respectively. We say system A converges faster than system B if there exists a time t $\in [0 < \tau < \infty]$, such that $\|x_i^A(\tau) - \delta_{i0}^*\|_2^A \leq \|x_i^B(\tau) - \delta_{i0}^*\|_2^B$ $\forall \quad i \in \{1, n\}$ and $\tau \in [t, \infty)$.This we will define as **Type-1 convergence**. In other words, system A will achieve faster **Type-1 convergence** as compared to system B.

Definition 2 : Assume there exist two systems A and B which satisfies the generic differential equations $\dot{x} = -Ax + \delta$ and $\dot{x} = -Bx + \delta$ respectively. We say A converges faster than B if there exists a time t $\in [0 < \tau < \infty]$, such that $\|x_i^A(\tau) - \delta_a^*\|_2^A \leq \|x_i^B(\tau) - \delta_b^*\|_2^B$ $\forall \quad i \in \{1, n\}$ and $\tau \in [t, \infty)$, where δ_a^* and $\delta_b *$ happens to be the steady state positions of the respective systems. This we will define as **Type-2 convergence**. In other words, system A will achieve faster **Type-2 convergence** as compared to system B.

Based on the definitions, we are ready to analyze the dependence of agent connectivity with its distortion time. Since our proposed attack does not destabilize the system,

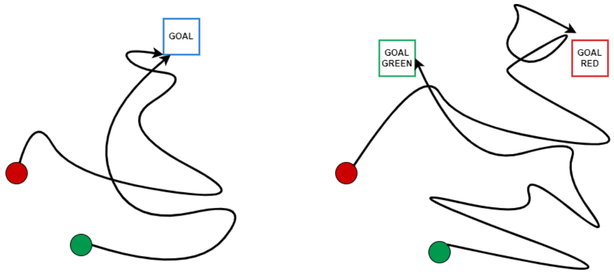

Fig. 4. Visual representation of Type 1 and Type 2 convergence

we expect the agents to converge at some point. This point of convergence will be different from that of the non-attacked case. Here, it is essential to note that **distortion time** can be taken as equivalent to the time taken for **Type 2 convergence**. We first need to study the parameters that directly influence the convergence time to establish the dependence. We begin our analysis with 'n' agents. The interaction topology in between those agents can be modelled in form of a graph $G = \{V, \epsilon\}$, such that $V \in \{1, n\}$, $\{E(i, j) = 1, j \in N_i\}$ and L be the Laplacian matrix of the graph G.

Theorem 2: Considering the attack scales as $\alpha_i \geq 0$, the second smallest eigenvalue (μ_2) of the compromised system (6) is bounded as follows :

$$\underline{\alpha_i}\left([\bar{L}]_{ii}\sqrt{[\bar{L}]_{ii}^{\,2} - i_c(G)^2}\right) \leq \mu_2 \leq \left(\sum_i \alpha_i\right)\lambda_n$$

Where, $\underline{\alpha_i} > 0$, $[\bar{L}]_{ii} \geq 0$, $i_c(G)$ is the minimum scale value, the maximum diagonal value of the Laplacian matrix ($[L]$), and the iso-perimetric number [29] corresponding to the weighted, directed graph G (underlying topology) respectively.

Proof: The lower bound of Theorem 2 is proved in [30]. For the upper bound, we can note that $\sum_i \mu_i = \sum_i [L]_{ii}\alpha_i$. Using Cauchy-Schwartz inequality, we can say $|\sum_i \mu_i| = |\sum_i [L]_{ii}\alpha_i|$. Expanding the R.H.S, we get $|\sum_i \mu_i| \leq \sqrt{(\sum_i \alpha_i^2)}\ \sqrt{(\sum_i [L]_{ii}^2)}$. Since, $\|x\|_2 \leq \|x\|_1$, $|\sum_i \mu_i| \leq (\sum_i \alpha_i)\cdot(\sum_i [L]_{ii})$. Since, trace($L$) = $\sum_i [L]_{ii} = \sum_i \lambda_i \leq (n-1)\lambda_n$, we can write $|\sum_i \mu_i| \leq (\sum_i \alpha_i)\cdot(n-1)\,\lambda_n$. From Theorem 1, we know $\mu_i \geq 0$, hence, $(n-1)\mu_2 \leq \sum_i \alpha_i\cdot(n-1)\,\lambda_n$. Thus, we found an upper bound on the second-smallest eigenvalue of the compromised system.

However, in un-directed weighted graphs, the results are somewhat simpler. Under the influence of the scaled attacks, the Laplacian undergoes structural changes, as shown in (3). As a result, the second smallest eigenvalue of the modified Laplacian needs to be analyzed[28]. In the un-directed case, the severity of the attack can be based on the distortion time. We need to understand that, attack severity is directly proportional to the time of distortion. Alternatively, for two attacks $\mathcal{A}$ and $\mathcal{B}$, the one with faster Type-2 convergence can be considered more severe. Corollary 2 states how the second smallest eigenvalues of the attacked and the non-attacked case in an undirected graph are related.

Corollary 1: As per the proposed formation control protocol, for time $\tau \in [t,\infty]$: $\|x_i(\tau) - \delta_{i0}^*\|_2^{FC} \leq \|x_i(\tau) - \delta_{i0}^*\|_2^{NFC}$ $\forall\, \mathrm{x}_i^{FC}(0) = x_i^{NFC}(0)$.

Proof : From the prior analysis we get:

$$x_i = \delta_{i0}^* + c_{1i}e^{-\lambda_1 t} + c_{2i}e^{-\lambda_2 t} + c_{3i}e^{-\lambda_3 t} + ...\ c_{ni}e^{-\lambda_n t}$$

$$x_i \leq \delta_{i0}^* + \ (|c_{1i}| + |c_{2i}| + |c_{3i}| + |c_{4i}|\ ...\ |c_{ni}|)\,e^{-\lambda_{n-1} t}$$

Thus, we see there a bound is being imposed by the second smallest eigenvalue i.e λ_{n-1}. In order to prove **Corollary 1**, we need to show the second smallest eigenvalue of a given graph G is largest in a fully-connected topology. Typically it is λ_{n-1} the eigenvalue that gives the most information about the graph (like expanding properties, connectivity, iso-perimetric properties, etc). Fiedler[21] in his classic paper showed that $\lambda_{n-1} \leq \frac{n}{n-1}\min\{d(v)\ : v \in V\}.\forall\, n > 1$ and $n - k > 0$ $\frac{n(n-k)}{n-1} > 0$. It also implies $\frac{n(n-k)}{n-1} < n$. Now, for all fully connected simple graph $d(v) = n - 1$ and for all other simple topology $\min\{d(v) : v \in V\} < n - 1$. This implies, $\lambda_{n-1}^{NFC} < n$. NFC stands for **non-fully connected case**.Since, $\lambda_{n-1}^{FC} = n$ we can claim $\lambda_{n-1}^{NFC} < \lambda_{n-1}^{FC}$. FC stands for **fully connected case**. So, for time $\tau \in [t,\infty]$: $\|x_i(\tau) - \delta_{i0}^*\|_2^{FC} \leq \|x_i(\tau) - \delta_{i0}^*\|_2^{NFC}$ $\forall\, \mathrm{x}_i^{FC}(0) = x_i^{NFC}(0)$. Thus, we can say a fully-connected topology will converge faster than all other non-fully connected case.

Corollary 2 : In the scaled attack setup, we can write the following two statements (**S1** and **S2**) :

S1:As per the attack mechanism on the proposed formation control protocol,$\|x_i^{LD_{i\alpha}}(\tau) - \delta_a^*\|_2^{LD_{i\alpha}} \leq \|x_i^L(\tau) - \delta_b^*\|_2^L$ $\forall$ $i \in \{1, n\}$ and $\tau \in [t, \infty)$.

S2 : The second smallest eigenvalue of the attacked system, $\mu_{n-1} \geq \min(\alpha_i).\lambda_{n-1}$. Where, α_i is the scaled attack parameter and λ_{n-1} is the second smallest eigenvalue of the non-attacked system.

Proof : System A converges faster than system B, if the second smallest eigenvalue of A is larger than that of B, where, A and B are graph Laplacians. In our case we have matrices namely L and $LD_{i\alpha}$.
Let μ_i be the eigenvalues of LS^{-1}. Then (λ_i, μ_i, s_i) obey the multiplicative version of Horn's inequalities [31]. The most basic of these, if $\lambda_1 \geq \lambda_2 ... \geq \lambda_n$and $s_1^{-1} \geq ... \geq s_n^{-1}$and $\mu_1 \geq \mu_2 \geq ... \geq \mu_n$ we can say $\mu_{i+j-1} \leq \lambda_i s_j^{-1}$ and $\mu_{i+j-n} \geq \lambda_i s_j^{-1}$.In our problem $S^{-1} = D_{i\alpha}$, and Spectrum($D_{i\alpha}$) $\in \{\alpha_1, \alpha_2, ..., \alpha_n\}$. Now, we can write $\mu_{n-1} \geq \lambda_{n-1}s_n^{-1}$ or $\mu_{n-1} \geq \lambda_{n-1}min(s_i^{-1})$ and eventually,$\mu_{n-1} \geq \lambda_{n-1}min(\alpha_i)$.Thus, we can claim that the second smallest eigenvalue of our compromised matrix is greater than or equal to that of the non-compromised case.

And, thus by the second part of Type-2 Convergence, Corollary 1 we can say that $\|x_i^{LD_{ik}}(\tau) - \delta_a^*\|_2^{LD_{ik}} \leq \|x_i^L(\tau) - \delta_b^*\|_2^L$ $\forall$ $i \in \{1, n\}$ **and** $\tau \in [t, \infty)$**.** Since, by **Theorem 2** we can say that $\lambda_{n-1}^{FC} \geq \lambda_{n-1}^{NFC}$, we can also extend the argument by saying $\mu_{n-1}^{FC} \geq \lambda_{n-1}^{FC} \geq \lambda_{n-1}^{NFC}$. And, $\mu_{n-1}^{NFC} \geq \lambda_{n-1}^{NFC}$. **This gives us an interesting interlacing theorem**.

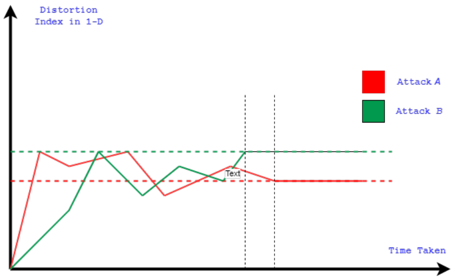

Fig. 5. In this example, in an un-directed weighted graph G, B achieves faster Type-2 convergence than A, we can say attack B is more severe as compared to attack A

VI. Attack analysis: Extent of Distortion

Considering the intruder has complete knowledge about the underlying topology and the desired formation, we assume the intruder can manipulate all the 'n' agents. We classify these forms of attacks as **'n-set-scaled attacks'**. Apart from that, we put up an additional constraint on the scales the intruder can use on those attacked agents. Given Δ_a,Δ_b $\in$ $\mathbb{R}^+$, the attacker can choose 'M' scale values distributed between [Δ_a,Δ_b]. In other words, we can say that $\alpha_i \in [\Delta_a,\Delta_b]$ such that $\Delta_b > \Delta_a > 0$.The distortion can be measured as the sum of net deviations from the desired values. Given δ_{ji} be the desired relative displacement between the j^{th} and the i^{th} agents and $\tilde{x}_i$ be the final position of the i^{th} agent under the influence of the scaled attack. $\|.\|_2$ represents the Euclidean norm. The measure of distortion error is defined as follows :

$$\epsilon_{rel} = \sum_i \sum_j \|\tilde{x}_i - \tilde{x}_j - \delta_{ij}\|_2^2 \tag{8}$$

Now, the above expression can be rewritten as follows.

$$\epsilon_{rel} = \sum_l \sum_i \sum_j (\tilde{x}_i^{(l)} - \tilde{x}_j^{(l)} - \delta_{ij}^{(l)})^2$$

$$\epsilon_{rel} = \sum_l {p^{(l)}}^T Q p^{(l)}$$

Where, $Q \in \mathbb{R}^{n\times n}$ and takes the special form as in (8), and $p^{(l)} = [\tilde{x}_1^{(l)} - \delta_1^{(l)} \;\; \tilde{x}_2^{(l)} - \delta_2^{(l)} \;\; ... \;\; \tilde{x}_n^{(l)} - \delta_n^{(l)}]$, and $p^{(l)} = \tilde{x}^{(l)} - \delta^{(l)}$.

$$Q = \begin{pmatrix} n-1 & -1 & -1 & -1 & ... & -1 \\ -1 & n-1 & -1 & -1 & ... & -1 \\ \vdots & \vdots & \vdots & \vdots & ... & \vdots \\ -1 & -1 & -1 & -1 & ... & n-1 \end{pmatrix} \tag{9}$$

From **Lemma 2**, we get the steady state value of the j^{th} component of the i^{th} agent.

$$\tilde{x}_i^{(j)} = \beta \left[\frac{\tilde{w}^T}{\alpha_i} (\alpha \odot x(0)^{(j)}) + \frac{\tilde{w}^T}{\alpha_i} K_i^j \right]$$

A. Optimal scale values for maximization of ϵ_{rel}

We will present our analysis for a sub-problem ($l = 1$) involving the first components of our agents. Given the axis are de-coupled, our analysis can easily be extended for the multi-dimensional case. The first component of the i^{th} agent is as follows :

$$\tilde{x}_i^{(1)} = \beta \left[\frac{\tilde{w}^T}{\alpha_i} (\alpha \odot x(0)^{(j)}) + \frac{\tilde{w}^T}{\alpha_i} (\mathbb{1}_n \delta_i^{(1)} - \delta^{(1)}) \right]$$

With the usual notations as stated above,

$$\tilde{x}_i^{(1)} = \beta \left[\frac{\bar{C}}{\alpha_i} + \frac{\sum_j \frac{\gamma_j^{(i)}}{\alpha_j}}{\alpha_i} \right]$$

Now, if we stack the components, we get the following:

$$x^{(1)} = \beta \Lambda \left[\bar{C} \mathbb{1}_n + \bar{B} \frac{1}{\alpha} \right] \tag{10}$$

Where,

$$\Lambda = \begin{pmatrix} \frac{1}{\alpha_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\alpha_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{\alpha_n} \end{pmatrix} \bar{B} = \begin{pmatrix} \gamma_1^{(1)} & \gamma_2^{(1)} & \dots & \gamma_n^{(1)} \\ \gamma_1^{(2)} & \gamma_2^{(2)} & \dots & \gamma_n^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_1^{(n)} & \gamma_2^{(n)} & \dots & \gamma_n^{(n)} \end{pmatrix}$$

The optimization problem can be stated as follows :

$$\begin{aligned} \max_{\alpha} \quad & 2 \cdot {p^{(1)}}^T Q p^{(1)} \\ \text{s.t.} \quad & \alpha \in [\Delta_a, \Delta_b] \odot \mathbb{1}_n \end{aligned}$$

This optimization problem is about maximizing the above function with respect to the hypercube expressed as constraints as stated above. Note that $p^{(l)}(\alpha)$ is written as $p^{(l)}$ throughout this paper. This problem is extremely hard to solve analytically, prompting us to present two of its special cases as an integral part of our analysis.

$Case_1$: Maximizing Distortion ϵ_{rel} under m-twin attack

The m-twin attack is a special case of scaled attack, where m ($< n$) nodes are attacked with scale values $\alpha_i = \bar{\alpha} \in [\Delta_a, \Delta_b]$ $\forall i \in \{1, m\}$. Unlike the general case stated in the section above, this problem is significantly easier to analyze. Since, there is only one scalar $\frac{1}{\bar{\alpha}}$, the optimization problem takes the special form :

$$\begin{aligned} \max_{\bar{\alpha}} \quad & \mathcal{F}\left(\frac{1}{\bar{\alpha}}\right) \\ \text{s.t.} \quad & \bar{\alpha} \in [\Delta_a, \Delta_b] \end{aligned}$$

Where $\mathcal{F}$ is a fractional polynomial defined on $\bar{\alpha}$. Our next step is to compute the extreme points by formulating the Lagrangian :

$$\mathcal{L} = \mathcal{F}\left(\frac{1}{\bar{\alpha}}\right) + \lambda_1 (\bar{\alpha} - \Delta_a) + \mu_1 (\bar{\alpha} - \Delta_b)$$

For λ or $\mu \neq 0$, from the K.K.T conditions, the extreme points are Δ_a or Δ_b. When both λ and μ equal to zero, the optimization problem defined in $Case_1$ reduces to the unconstrained problem. To proceed we begin with :

$$x_i^{(1)} = \left[w^T x(0)^{(1)} + \sum_j^{j \in \mathcal{A}} \frac{1}{\bar{\alpha}} \gamma_j^{(1)} + \sum_j^{j \in \mathcal{A}^C} \gamma_j^{(1)} \right] \frac{\beta}{\alpha_i}$$

Here, $j \in \mathcal{A}$, $j \in \mathcal{A}^C$ refers to the attacked nodes and the normal nodes respectively. Taking $a_i = w^T x(0)^{(1)} + \sum_j^{j \in \mathcal{A}^C} \gamma_j^{(1)}$, $\bar{b}_i = \sum_j^{j \in \mathcal{A}} \frac{1}{\bar{\alpha}} \gamma_j^{(1)}$ and $\beta = \frac{\bar{\alpha}}{m+(1-m)\bar{\alpha}}$ we can write the ϵ_{rel} as :

$$\epsilon_{rel} = \beta^2 \sum_i \sum_j \left[(a_i + \bar{b_i} \frac{1}{\bar{\alpha}}) \frac{1}{\alpha_i} - (a_j + \bar{b_j} \frac{1}{\bar{\alpha}}) \frac{1}{\alpha_j} - \frac{\delta_{ij}}{\beta} \right]^2$$

Considering $\frac{1}{\alpha_i} = \frac{1}{\bar{\alpha}}$ $(i \in \mathcal{A})$,$\frac{1}{\alpha_i} = 1(i \in \mathcal{A}^C)$, and $\mathcal{F}\left(\frac{1}{\bar{\alpha}}\right)$ is a polynomial function of $\frac{1}{\bar{\alpha}}$.Hence, we can write the expression as ϵ_{rel} as :

$$\epsilon_{rel} = \beta^2 \mathcal{F}\left(\frac{1}{\bar{\alpha}}\right)$$

$$\frac{\partial \epsilon_{rel}}{\partial \bar{\alpha}} = \mathcal{F}\left(\frac{1}{\bar{\alpha}}\right) = 0 \, ; \; \frac{\partial^2 \epsilon_{rel}}{\partial^2 \bar{\alpha}} = \dot{\mathcal{F}}\left(\frac{1}{\bar{\alpha}^*}\right) < 0 \qquad (11)$$

In order to find the maximum ϵ_{rel} for $\bar{\alpha} \in [\Delta_a, \Delta_b]$ we take the derivative w.r.t $\bar{\alpha}$ and we will end up with polynomial $\mathcal{F}\left(\frac{1}{\bar{\alpha}}\right)$ in (10). Hence, $\bar{\alpha}^* \in [\Delta_a, \Delta_b]$ satisfying the polynomial $\mathcal{F}\left(\frac{1}{\bar{\alpha}^*}\right) = 0$ and $\dot{\mathcal{F}}\left(\frac{1}{\bar{\alpha}^*}\right) < 0$ is the optimal scale value that maximizes ϵ_{rel}.

B. *$Case_2$: Maximizing Distortion when δ_{ij} = 0*

When we put δ_{ij} = 0, equation (10) reduces to the following :

$$x^{(1)} = \beta \Lambda \left[\bar{C} \mathbb{1}_n \right] \qquad (12)$$

Since, δ_{ij} is taken as 0, the optimization problem becomes equivalent to

$$\begin{aligned} \max_{\alpha} \quad & 2 \cdot {x^{(1)}}^T Q x^{(1)} \\ \text{s.t.} \quad & \alpha \in [\Delta_a, \Delta_b] \odot \mathbb{1}_n \end{aligned}$$

The Lagrangian in this case will be :

$$\mathcal{L} = 2\bar{C}^2 \beta^2 \mathbb{1}_n^T \Lambda^T Q \Lambda \mathbb{1}_n + \sum_i \lambda_i (\alpha_i - \Delta_a) + \sum_i \mu_i (\alpha_i - \Delta_b)$$

From the K.K.T conditions, we get 2^n possible extreme points given λ_i and μ_i $\forall i, j \in \{1, n\}$ are active. When they are not active, the unconstrained version gives us :

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 4\bar{C}^2 \beta^2 Q \Lambda \mathbb{1}_n + 2\bar{C}^2 \mathbb{1}_n^T \Lambda^T Q \Lambda \mathbb{1}_n \beta \dot{\beta} = 0$$

To find the global maximum, we have to check $2^n + n$ points, making the problem computationally challenging.

VII. Optimal topology for minimizing ϵ_{rel} in a compromised system

Our next goal is to find the optimal topology (G^*) that minimizes the distortion (ϵ_{rel}) under the influence of the attack defined in (3). In order to find that, we assume the attack scales are being provided, and the intent is to find w^* and eventually L^*, which serves the purpose. Keeping that in mind, we can formulate our optimization problem as follows:

$$\begin{aligned} \min_{w} \quad & 2 \cdot \sum_l {p^{(l)}}^T Q p^{(l)} \\ \text{s.t.} \quad & 0 \le w_i \le 1 \;\; \forall i \in \{1, n\}, \\ & \sum_{i=1}^{n} w_i = 1 \end{aligned}$$

A. *Sub-optimal Topology for Minimizing Distortion under fixed scales*

As encountered previously, solving the family of multivariate polynomial is hard in general. But, certain acceptable sub-optimal strategies can also turn out to be useful, For example, in this case, we assume the attacker has already chosen the scales (α_i). We need to find the underlying topology most vulnerable to the designed attack. Again we will present our analysis for a sub-problem ($l = 1$) involving the first components of our agents. A similar approach can be generalized for multi-dimensional agents since the components were decoupled in our problem formulation. We start with the steady state of the first component of the i^{th} agent :

$$\tilde{x}_i^{(1)} = \beta \left[\frac{\tilde{w}^T}{\alpha_i} (\alpha \odot x(0)^{(j)}) + \frac{\tilde{w}^T}{\alpha_i} K_i^1 \right]$$

Given α_i's are already chosen, the above expression can clearly be written as this :

$$\tilde{x}_i^{(1)} = \beta \left[\sum_l \gamma_l^i \cdot w_l \right]$$

Taking the steady state difference between the i^{th} agent and the j^{th} agent we get :

$$\begin{aligned} \tilde{x}_i^{(1)} - \tilde{x}_j^{(1)} &= \beta \left[\sum_l \gamma_l^i \cdot w_l - \sum_l \gamma_l^j \cdot w_l \right] \\ &= \beta \left[\sum_l \gamma_l^{ij} \cdot w_l \right] \;\; ; \;\; \gamma_l^{ij} = \gamma_l^i - \gamma_l^j \end{aligned}$$

Following the previous footsteps, we can formulate the sum of the squared difference as follows:

$$\sum_i \sum_j \left(\tilde{x}_i^{(1)} - \tilde{x}_j^{(1)} - \delta_{ij}^{(1)} \right)^2 = \sum_i \sum_j \beta^2 \left[\sum_l \gamma_l^{ij} \cdot w_l - \frac{\delta_{ij}^{(1)}}{\beta} \right]^2$$

If we consider $\gamma_{n+1}^{ij} = -\frac{\delta_{ij}^{(1)}}{\beta}$ and $w_{n+1} = 1$:

$$\sum_i \sum_j \left(\tilde{x}_i^{(1)} - \tilde{x}_j^{(1)} - \delta_{ij}^{(1)} \right)^2 = \sum_i \sum_j \beta^2 \left[\sum_{l=1}^{n+1} \gamma_l^{ij} \cdot w_l \right]^2$$

It can be observed that the summation $\sum_l \gamma_l^{ij} \cdot w_l$ is basically the dot-product $({\gamma^{ij}}^T w)$ expressed as the product of the Euclidean norms of the individual vectors (γ^{ij}, w) and the angle between them $(\theta_{\gamma^{ij} w})$. Hence, we get the following expression :

$$\sum_i \sum_j \left(\tilde{x}_i^{(1)} - \tilde{x}_j^{(1)} - \delta_{ij}^{(1)}\right)^2 = \sum_i \sum_j \beta^2 \cdot \|\gamma^{ij}\|_2^2 \cdot \|w\|_2^2 cos^2(\theta_{\gamma^{ij} w})$$

The 2-norm, and the 1- norm are expressed by an inequality $\|x\|_2 \leq \|x\|_1 \leq \sqrt(n)\|x\|_2$. Hence, we get this strong upper and lower bound of our principal objective function.

$$\sum_i \sum_j \frac{\beta^2}{n} \cdot \|\gamma^{ij}\|_2^2 \cdot \|w\|_1^2 cos^2(\theta_{\gamma^{ij} w}) \leq \sum_i \sum_j \left(\tilde{x}_i^{(1)} - \tilde{x}_j^{(1)} - \delta_{ij}^{(1)}\right)^2 \\ \leq \sum_i \sum_j \beta^2 \cdot \|\gamma^{ij}\|_2^2 \cdot \|w\|_1^2 cos^2(\theta_{\gamma^{ij} w})$$

Since, $\sum_i w_i$ **=** $\|w\|_1$ **= 1 , and** $\sum_{i=1}^{n+1} w_i = 2$ **we can write** :

$$2\sum_i \sum_j \frac{\beta^2}{n} \cdot \|\gamma^{ij}\|_2^2 \cdot cos^2(\theta_{\gamma^{ij} w}) \leq \sum_i \sum_j \left(\tilde{x}_i^{(1)} - \tilde{x}_j^{(1)} - \delta_{ij}^{(1)}\right)^2 \\ \leq 2\sum_i \sum_j \beta^2 \cdot \|\gamma^{ij}\|_2^2 \cdot cos^2(\theta_{\gamma^{ij} w})$$

Writing $K_{\gamma^{ij}} = \|\gamma^{ij}\|_2^2 \cdot \beta^2$ we can express the previous expression in a compact form. Let $\bar{K}_{\gamma^{ij}}$ and $\underline{K}_{\gamma^{ij}}$ be the minimum and maximum value of $K_{\gamma^{ij}}$ respectively. Finally, we found two strong bounds of the distortion metric, encapsulating the effect of underlying topology in determining the severity of the attack under a chosen set of scales :

$$2\underline{K}_{\gamma^{ij}} \cdot \sum_i \sum_j cos^2(\theta_{\gamma^{ij} w}) \leq n \cdot \epsilon_{rel} \\ \leq 2n \cdot \bar{K}_{\gamma^{ij}} \cdot \sum_i \sum_j cos^2(\theta_{\gamma^{ij} w}) \tag{13}$$

Assuming a w^* exists that minimizes the upper bound and lower bound in (10), we can say there exists a Laplacian L^* such that

$$w^{*T} L^* = 0 \tag{14}$$

The L^* defined in (11) gives us the underlying graph (G^*), which minimizes the distortion (ϵ_{rel}) under the influence of the given scaled attack.

B. Sub-optimal strategies to find w^*

In this section, we will present one sub-optimal strategy to find w^*, such that the upper and lower bound in (13) is minimized. In that way, we can claim, the global minima should lie in that given range. Let $\gamma = \{(i,j) \in \{1,n\}|\gamma_{ij}\}$ be partitioned into 2-nearest neighbor sets (γ_1 and γ_2) such that $\gamma = \gamma_1 \oplus \gamma_2$. Considering,$\bar{\gamma_1}$ and $\bar{\gamma_2}$ as the average of the two chosen sets, we just need to choose $w^* = \mathcal{K} \cdot (\bar{\gamma_1} \times \bar{\gamma_2})$, so that w^* is perpendicular to this generated 2-nearest set vectors (γ_1, γ_2). This method can be further refined by dropping the outliers while computing γ_1 and γ_2.

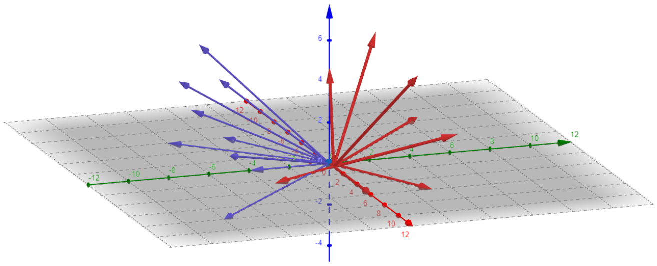

Fig. 6. The red vectors and the blue vectors respectively form the 2 components of K-nearest neighbor in 3 dimensions. The average of the red vectors(γ_1), and the blue vectors (γ_2) will help us choose w^* sub-optimally.

VIII. Lower bound on the left-dominant eigenvectors of a Graph and its Sub-graph

A. Relative distortion error and its relation to connectivity

Connectivity of a graph $G = (V, \mathcal{E})$ is defined as the number of edges the graph has (in other words, $|\mathcal{E}|$). We say graph $G_1 = (V, \mathcal{E}_1)$ is more connected as compared to $G_2 = (V, \mathcal{E}_2)$ if and only if $|\mathcal{E}_1| > |\mathcal{E}_2|$.We change our formal measure of distortion to introduce the relative distortion error. $\|.\|_2^H$ and $\|.\|_2^L$ represents Euclidean norm in higher and lower connectivity, respectively. The modified relative distortion error is as follows :

$$\epsilon_{rel} = \sum_i \sum_j (\|\tilde{x}_i^H - \tilde{x}_j^H - \delta_{ij}\|_2^H - \|\tilde{x}_i^L - \tilde{x}_j^L - \delta_{ij}\|_2^L) \tag{15}$$

$$\epsilon_{rel} = \epsilon_{RDE}^H - \epsilon_{RDE}^L \tag{16}$$

B. Relative distortion error in undirected, weighted topology

We define higher and lower connectivity in an undirected, weighted topology as in the previous section. We can start by computing ϵ_{RDE}^H and ϵ_{RDE}^L for the weighted, undirected case. Since undirected topologies are always weight balanced, we can say $\tilde{w} = \mathbb{1}_n$ for all laplacians. Hence,

$\epsilon_{RDE}^H = \epsilon_{RDE}^L$ and $\epsilon_{rel} = \epsilon_{RDE}^H - \epsilon_{RDE}^L = 0$
Thus, distortion in un-directed, weighted topology stays unchanged in any underlying network topology. This is because in weighted, undirected networks composed of fixed vertices, the eigenvector corresponding to 0 eigenvalue remains unchanged.
Theorem 3 : Given two graph laplacians L_h and L_l , the euclidean distance between the left dominant eigenvectors($\tilde{w}_H, \tilde{w}_L$) satisfying $\mathbb{1}_n^T \tilde{w}_H (\succcurlyeq 0) = \mathbb{1}_n^T \tilde{w}_L (\succcurlyeq 0) = 1$ and $\tilde{w}_H^T L_h = \tilde{w}_L^T L_l = \mathbb{0}_n^T$ is bounded by $\|\tilde{w}^H - \tilde{w}^L\|_2 \geq \frac{1}{r} \sum_{n=1}^r \frac{|\sum_{i=1}^m \tilde{w}_i \sigma_{in}|}{\|C_n\|_2}$.

Proof of Theorem 3: In the above theorem, $\|C_n\|_2$ is the euclidean norm of the n^{th} column of G and σ_{ij} is the change in the ij^{th} entry of G and H. Let $G = \{V, \mathcal{E}_1\}$, $H = \{V, \mathcal{E}_2\}$ be a weighted digraph and it's subgraph respectively. Mathematically, $G \succcurlyeq H$ also holds the same meaning. For both G and H, we can write the graph Laplacian as L_G and L_H respectively.Let the left dominant eigenvectors(associated with 0 eigenvalues) associated with L_G and L_H be $\tilde{w}_H$ and $\tilde{w}_L$ respectively.

$$\tilde{w}_H^T L_h = \tilde{w}_H^T L_h = \mathbb{0}_n^T$$

Where,

$$\tilde{w}_H^T = [w_1 \;\; w_2 \;\; w_3 \;\; ... \;\; w_n] \;\; , \tilde{w}_L^T = [\tilde{w}_1 \;\; \tilde{w}_2 \;\; \tilde{w}_3 \;\; ... \;\; \tilde{w}_n]$$

Similarly,

$$L_h = [C_1^h \;\; C_2^h \;\; C_3^h \;\; ... \;\; C_n^h] \;\; , L_l = [C_1^l \;\; ... \;\; C_r^l \;\; C_{r+1}^h \;\; ... \;\; C_n^h]$$

We assumed that $(n-r)$ columns in L_l is same that of L_h, and r columns have changed as per the change in weights that resulted due to removal of edges. In other words, $L_g \ngeq L_l$.
Here, $w_H^T C_1^h = \tilde{w}_L^T C_1^l$ and so on. If we write $C_1^h = [a_{11} \;\; a_{21} \;\; a_{31} \;\; ... \;\; a_{n1}]$ and $C_1^l = [\tilde{a}_{11} \;\; \tilde{a}_{21} \;\; \tilde{a}_{31} \;\; ... \;\; \tilde{a}_{n1}]$. Expanding the previous expression, $\sum_i w_i a_{i1}$=$\sum_i \tilde{w}_i \tilde{a}_{i1}$. Assuming, there are 'm' changed entries in C_1^l which is a result of edge-removal from G. Further, the change in the ij^{th} entry in L_l is denoted by σ_{ij}. Hence, for the first column we write $\sum_i \tilde{w}_i \sigma_{i1} = \sum_i (w_i - \tilde{w}_i) a_{i1}$. Now taking modulus on both sides we get, $|\sum_i \tilde{w}_i \sigma_{i1}|$=$|\sum_i (w_i - \tilde{w}_i) a_{i1}|$. By using Cauchy-Schwartz Inequality on both the sides, $|\sum_i \tilde{w}_i \sigma_{i1}| \leq \; \|\tilde{w}^H - \tilde{w}^L\|.\|C_1\|_2$.Hence,$\|\tilde{w}^H - \tilde{w}^L\|_2 \geq \frac{|\sum_i \tilde{w}_i \sigma_{i1}|}{\|C_1\|_2}$. Proceeding similarly, we can make this lower bound more refined as follows :

$$\|\tilde{w}^H - \tilde{w}^L\|_2 \geq \frac{1}{r} \sum_{n=1}^{r} \frac{|\sum_{i=1}^{m} \tilde{w}_i \sigma_{in}|}{\|C_n\|_2}$$

IX. Illustrative Example

For $Case_1$, we present an illustrative example as follows. For the given network below, x(0) = [0.356 0.384 0.4123 0.4417 0.223 -0.6813] the normalized left-dominant eigenvector $\tilde{w}$ = [0.13 0 0.34 0.2316 0.235 0.241]the fractional polynomial be $\mathcal{F} = \frac{1.4996\bar{\alpha}^4 - 8.36\bar{\alpha}^2 + 6.32\bar{\alpha} + 6}{3.3\bar{\alpha} + 6.7}$ where $\bar{\alpha} \in [0.08, 2]$. We see the value of $\bar{\alpha}$ that maximizes this $\mathcal{F}$ is 0.195. And, ϵ_{rel} equals to 0.942.

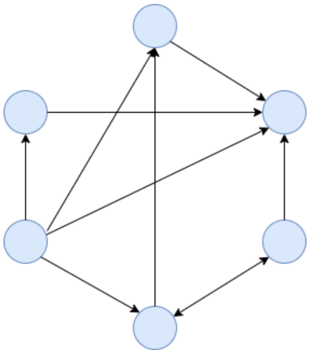

Fig. 7.

X. Conclusions

The paper introduces a scaled attack strategy aimed at distorting the formation schemes of agents within a multi-agent system. The attack scheme, while preserving the underlying network topology, manipulates the convergence time in a structured manner, presenting a disruptive influence upon the system dynamics. In the first part of the study, emphasis is placed on maintaining the invariant topology of the network. Through the proposed attack scheme, the convergence time is deliberately altered, introducing intentional delays to the previously predictable process. This orchestrated interference impacts the agents' ability to reach a synchronized state efficiently, effectively disrupting the system's desired convergence.In the subsequent section, the attack strategy sets its sights on identifying and exploiting scales to maximize the steady state distortion. Despite the inherent complexity of the original problem, the authors successfully tackle two simplified cases, offering valuable insights into potential approaches for addressing the broader challenge. By systematically exploring the relationship between scales and the resulting distortion, a general methodology for tackling the problem is elucidated.Furthermore, the paper delves into the prospect of searching for an optimal network topology capable of minimizing distortion for a given set of scales. The authors undertake a thorough exploration of this possibility, seeking to identify network configurations that mitigate the disruptive impact of the nodal attack strategy. By assessing the convergence properties of two similar networks, they also endeavor to establish upper bounds on these convergence times, providing valuable bounds for future analysis.In summary, the research presents a nodal attack strategy designed to disturb the formation schemes of agents within multi-agent systems. Through deliberate alterations to convergence time and the pursuit of scale-based distortions, the authors navigate the complexities of the problem, offering insights into potential mitigation strategies and upper bounds for convergence properties. This work contributes to the broader understanding of adversarial influences within multi-agent systems and sets the stage for further investigations into the optimization of network topologies to counter such attacks.

Acknowledgment

In the realm of my project endeavors, I find myself compelled to extend my heartfelt appreciation to the exceptional Dr. Vaibhav Katewa, my revered supervisor. Throughout this remarkable journey, his unwavering support and ceaseless motivation have been a beacon of light.Apart from that, I feel blessed to have crossed path with Sameeksha Singh in the second year of my stay at Indian Institute of Science. The past one year was possibly one of the toughest year of my life, in which a lot of salient decisions had to be made. Yet always, I was always at ease owing to her mere presence, which I believe has done wonders to my life. She has lent a patient ear to all my troubles, and her lively and vibrant spirit has kept me going at all times. I have no words to express my gratitude towards her.

References

[1] Y. Lee, J. Trevathan, I. Atkinson, and W. Read, “An intelligent agent system for managing heterogeneous sensors in dispersed and disparate wireless sensor network,” Int. J. Sens. Netw., vol. 27, no. 3, pp. 149–162, 2018.

[2] J. Gao et al., “SCADA communication and security issues,” Security Commun. Netw., vol. 7, no. 1, pp. 175–194, 2014

[3] Y. Cao, W. Yu, W. Ren, and G. Chen, “An overview of recent progress in the study of distributed multi-agent coordination,” IEEE Trans. Ind. Informat., vol. 9, no. 1, pp. 427–438, Feb. 2013.

[4] K.-K. Oh, M.-C. Park, and H.-S. Ahn, “A survey of multi-agent formation control,” Automatica, vol. 53, pp. 424–440, Mar. 2015.

[5] Y. Lee, J. Trevathan, I. Atkinson, and W. Read, “An intelligent agent system for managing heterogeneous sensors in dispersed and disparate wireless sensor network,” Int. J. Sens. Netw., vol. 27, no. 3, pp. 149–162, 2018.

[6] Y. Yang, Y. Xiao, and T. Li, “A survey of autonomous underwater vehicle formation: Performance, formation control, and communication capability,” IEEE Commun. Surveys Tuts., vol. 23, no. 2, pp. 815–841, 2nd Quart., 2021.

[7] I. Shames, A. M. H. Teixeira, H. Sandberg, and K. H. Johansson, “Distributed fault detection for interconnected second-order systems,” Automatica, vol. 47, no. 12, pp. 2757–2764, 2011.

[8] H. J. LeBlanc and X. D. Koutsoukos, “Consensus in networked multiagent systems with adversaries,” in Proc. 14th Int. Conf. Hybrid Syst. Comput. Control, Chicago, IL, USA, Apr. 2011, pp. 281–290.

[9] Z. Feng, G. Wen, and G. Hu, “Distributed secure coordinated control for multiagent systems under strategic attacks,” IEEE Trans. Cybern., vol. 47, no. 5, pp. 1273–1284, May 2017.

[10] M. Zhu and S. Martinez, “On the performance analysis of resilient networked control systems under replay attacks,” IEEE Trans. Autom. Control, vol. 59, no. 3, pp. 804–808, Mar. 2014.

[11] X.-M. Li, Q. Zhou, P. Li, H. Li, and R. Lu, “Event-triggered consensus control for multi-agent systems against false data-injection attacks,” IEEE Trans. Cybern., vol. 50, no. 5, pp. 1856–1866, May 2020.

[12] X.-M. Zhang, Q.-L. Han, X. Ge, and L. Ding, “Resilient control design based on a sampled-data model for a class of networked control systems under denial-of-service attacks,” IEEE Trans. Cybern., vol. 50, no. 8, pp. 3616–3626, Aug. 2020.

[13] C. Peng and H. Sun, “Switching-like event-triggered control for networked control systems under malicious denial of service attacks,” IEEE Trans. Autom. Control, vol. 65, no. 9, pp. 3094–3103, Sep. 2020.

[14] T.-Y. Zhang and D. Ye, “Distributed secure control against denial-ofservice attacks in cyber-physical systems based on K-connected communication topology,” IEEE Trans. Cybern., vol. 50, no. 7, pp. 3094–3103,Jul. 2020.

[15] M. Cao, F. Xiao, and L. Wang, “Event-based second-order consensus control for multi-agent systems via synchronous periodic event detection,” IEEE Trans. Autom. Control, vol. 60, no. 9, pp. 2452–2457, Sep. 2015.

[16] M. A. Lewis and K.-H. Tan, “High precision formation control of mobile robots using virtual structures,” Auton. Robots, vol. 4, no. 4, pp. 387–403, 1997.

[17] Z. Yan, X. Pan, Z. Yang, and L. Yue, “Formation control of leaderfollowing multi-UUVs with uncertain factors and time-varying delays,” IEEE Access, vol. 7, pp. 118792–118805, 2019.

[18] M. Conti, N. Dragoni, and V. Lesyk, “A survey of man in the middle attacks,” IEEE Commun. Surveys Tuts., vol. 18, no. 3, pp. 2027–2051, 3rd Quart., 2016.

[19] J. Deng, K. Meng, Y. Xiao, and R. Xu, “Implementation of DoS attack and mitigation strategies in IEEE 802.11b/g WLAN,” in Proc. Sens. Command Control Commun. Intell. (C3I) Technol. Homeland Security Homeland Defense IX, vol. 7666. Orlando, FL, USA, Apr. 2010, pp. 29–38

[20] X.-M. Zhang et al., “Networked control systems: A survey of trends and techniques,” IEEE/CAA J. Autom. Sinica, vol. 7, no. 1, pp. 1–17, Jan. 2020.

[21] M. Fiedler, “Algebraic connectivity of graphs,” Czechoslovak mathematical journal, vol. 23, no. 2, pp. 298-305, 1973.

[22] X. Shao and D. Ye, “Fuzzy adaptive event-triggered secure control for stochastic nonlinear high-order mass subject to DoS attacks and actuator faults,” IEEE Trans. Fuzzy Syst., early access, Oct. 5, 2020, doi: 10.1109/TFUZZ.2020.3028657.

[23] Y. Yang, Y. Xiao and T. Li, ”Attacks on Formation Control for Multiagent Systems,” in IEEE Transactions on Cybernetics, vol. 52, no. 12, pp. 12805-12817, Dec. 2022, doi: 10.1109/TCYB.2021.3089375.

[24] Q. Ali and S. Montenegro, ”Role of graphs for multi-agent systems and generalization of Euler’s Formula,” 2016 IEEE 8th International Conference on Intelligent Systems (IS), Sofia, Bulgaria, 2016, pp. 198-204, doi: 10.1109/IS.2016.7737421.

[25] Z. Feng, G. Wen, and G. Hu, “Distributed secure coordinated control for multiagent systems under strategic attacks,” IEEE Trans. Cybern., vol. 47, no. 5, pp. 1273–1284, May 2017.

[26] F. Bullo, ”Lectures on Networked Control Systems”, 6th edition pp. 127

[27] M. Pirani and S. Sundaram, ”Spectral properties of the grounded Laplacian matrix with applications to consensus in the presence of stubborn agents,” 2014 American Control Conference, Portland, OR, USA, 2014, pp. 2160-2165, doi: 10.1109/ACC.2014.6859421.

[28] Wei Ren and R. W. Beard, ”Consensus seeking in multiagent systems under dynamically changing interaction topologies,” in IEEE Transactions on Automatic Control, vol. 50, no. 5, pp. 655-661, May 2005, doi: 10.1109/TAC.2005.846556.

[29] S. Danda, A. Challa, B. S. Daya Sagar and L. Najman, ”Revisiting the Isoperimetric Graph Partitioning Problem,” in IEEE Access, vol. 7, pp. 50636-50649, 2019, doi: 10.1109/ACCESS.2019.2901094.

[30] Abraham Berman, Xiao-Dong Zhang ”Lower bounds for the eigenvalues of Laplacian matrices”