Identification of locally influential agents in self-organizing multi-agent systems

Kshitij Jerath[1] and Sean Brennan[2]

Abstract— **Current research methods directed towards measuring the influence of specific agents on the dynamics of a large-scale multi-agent system (MAS) rely largely on the notion of controllability of the full-order system, or on the comparison of agent dynamics with a user-defined macroscopic system property. However, it is known that several large-scale multi-agent systems tend to self-organize, and their dynamics often reside on a low-dimensional manifold. The proposed framework uses this fact to measure an agent's influence on the macroscopic dynamics. First, the minimum embedding dimension that can encapsulate the low-dimensional manifold associated with the self-organized dynamics is identified using a modification of the method of false neighbors. Second, the minimum embedding dimension is used to guide the Krylov subspace-based model order reduction of the system dynamics. Finally, an existing controllability-based metric is applied to the local reduced-order representation to measure an agent's influence on the self-organized dynamics. With this technique, one can identify regions of the state space where an agent has significant local influence on the dynamics of the self-organizing MAS. The proposed technique is demonstrated by applying it to the problem of vehicle cluster formation in traffic, a prototypical self-organizing system. As a result, it is now possible to identify regions of the roadway where an individual driver has the ability to influence the dynamics of a self-organized traffic jam.**

I. Introduction

In recent years, there has been significant interest in the complex behavior exhibited by large-scale multi-agent systems (MAS). The behavior of such systems, e.g. synchronization of coupled oscillators, flocking of birds etc., is referred to as *self-organization* because the system dynamics often exhibit spatio-temporal patterns that occur in the absence of external control [1]. However, by their very nature, it is difficult to control the global (or macroscopic) behavior of self-organizing MAS because (i) an external centralized mechanism to control the macroscopic dynamics of such a large-scale system would likely require significant control effort, and (ii) alternative de-centralized approaches would require controlling each agent in a large-scale MAS, which is often not a feasible option. However, macroscopic behavior of self-organizing systems may be *influenced* by controlling a small subset of agents.

Such a phenomenon can already be seen at play in several natural-engineered systems, where the self-organizing behavior of a naturally existing large-scale multi-agent system is affected by introduction of a 'small' set of artificially-engineered agents. Examples include the effects of vehicles equipped with adaptive cruise control (ACC) on formation of self-organized jams in a traffic flow consisting of human-driven vehicles [2], and the use of intra-cerebral implants to control seizures in the human brain [3]. These systems are of even greater interest because they have considerable societal impact – traffic congestion in the USA resulted in losses that exceeded USD 100 billion (in 2011 dollars), and epileptic seizures affect more than 50 million people worldwide [4].

Within a MAS, not all agents have the same influence on the system's macroscopic dynamics. The following work presents a methodology to choose the set of most influential agents with respect to the macroscopic behavior of the naturally existing large-scale MAS. Self-organized vehicle cluster formation is used as a prototypical example to demonstrate this methodology. The remainder of this paper is organized as follows: Section II discusses prior work on measures of influence and identification of influential agents in a large-scale system. Section III discusses the self-organized dynamics observed during vehicle cluster formation. Section IV outlines the use of the method of false neighbors to determine the minimum embedding dimension of the low-dimensional manifold on which the dynamics of the self-organizing MAS evolve. Section V describes the Krylov subspace-based method to obtain a local reduced-order representation of the low-dimensional manifold, and uses this information, along with existing controllability-based metrics of influence, to determine the ability of an agent to influence macroscopic dynamics. Section VI summarizes the key insights obtained from this work.

II. Literature review

In the past decade, several researchers have studied the influence of specific agents in networked multi-agent systems. Many such approaches use a combinatorial search strategy to determine the set of agents that maximize an output-based measure, i.e. a pre-specified macroscopic quantity of the network. Within the framework of pinning control [5], Porfiri and di Bernardo have proposed a technique to identify a set of pinned nodes or agents that lead to higher network synchronization strength [6]. Similarly, Patterson et al. have suggested the use of network coherence as the macroscopic quantity to optimize during the combinatorial search [7]. Recent works have also studied the influence of individual agents on macroscopic properties by drawing inspiration from other fields, such as the manipulability measure of influence proposed by Kawashima et al., which

[1]Kshitij Jerath is jointly with the Department of Mechanical and Nuclear Engineering, and the Department of Electrical Engineering.

[2]Sean Brennan is with the Department of Mechanical and Nuclear Engineering. Both authors are at The Pennsylvania State University, University Park, PA 16802, USA. (e-mail: kjerath@outlook.com, sbrennan@psu.edu)

draws inspiration from sensitivity-like measures developed for robotic arms [8]. Often, the optimized macroscopic quantity (or output-based measure) is selected based on the user's understanding of the system dynamics (such as studying synchronization strength in coupled oscillators, or network coherence in consensus problems). However, in systems where the underlying self-organized dynamics are not well understood, approaches that rely on output-based measures may have limited applicability.

Other researchers have approached the problem more directly by attempting to develop measures of influence for individual agents. The influence metrics developed in recent times have largely relied on controllability-based approaches, such as using the determinant, minimum eigenvalue, or other properties of the controllability grammian [9][10][11]. These works follow a rich tradition of analyzing controllability, and measuring the difference between controllable and uncontrollable systems, such as presented in the early works of Eising [12], Hamdan [13], and Viswanathan et al. [14]. Additionally, the spirit of these works is embodied in some early research performed by Lim to determine the optimal location of actuators and sensors in large-scale flexible space structures [15]. In this work, Lim performed a combinatorial search to determine optimal locations of the actuators and sensors that improved the controllability and observability of the flexible space structure.

These prior works illustrate the need for a methodology to assess agent influence in *self-organizing* systems. The following work describes such a systematic methodology while using the formation of vehicle clusters as a prototypical example of a self-organizing multi-agent system.

III. Vehicle cluster formation: a prototypical self-organizing system

Self-organized vehicle clusters, also known as 'phantom jams' and 'stop-and-go waves', are known to form in medium-to-high density traffic. Such vehicle cluster formation has been observed in traffic and experiments, and can be replicated with commonly-used models of driver behavior. Over the years, several car-following models have been proposed for driver behavior, and most have a general form given by:

$$\ddot{x}_i(t+\tau) = f(\gamma_i, x_i, x_{i-1}, \dot{x}_i, \dot{x}_{i-1}) \tag{1}$$

where $\ddot{x}_i$ denotes the acceleration of the following vehicle, τ denotes the delay in the response of the following vehicle, γ_i denotes the driver sensitivity of the following vehicle, x_i the position of the following vehicle, x_{i-1} denotes the position of the followed vehicle, and $\dot{x}_i$ and $\dot{x}_{i-1}$ denote the velocities of the following and followed vehicles, respectively. For this study, driver behavior will be described by the intelligent driver model (IDM) because it is one of the few models that can mimic car-following dynamics without requiring an explicit delay term [16]. The absence of an explicit delay term is advantageous because it helps avoid modeling the system via delay differential equations and significantly simplifies the ensuing controllability-based analysis.

A. Intelligent Driver Model

The intelligent driver model is expressed as:

$$\dot{x}_i(t) = v_i \tag{2}$$

$$\begin{aligned}\ddot{x}_i(t) &= f(x_i, x_{i-1}, \dot{x}_i, \dot{x}_{i-1}, a_i, b_i) \\ &= a_i \left\{ 1 - \left(\frac{v_i}{v_0}\right)^{\delta} - \left(\frac{s(v_i, \Delta v_i)}{s_i}\right)^2 \right\}\end{aligned} \tag{3}$$

where, $s_i = x_i - x_{i-1}$ denotes the spacing between the following and followed vehicles, $\Delta v_i = v_i - v_{i-1} = \dot{x}_i - \dot{x}_{i-1}$ denotes the relative velocity between the following and followed vehicles, $s(v_i, \Delta v_i) = s_0 + v_i T + v_i \Delta v_i / (2\sqrt{a_i b_i})$, a_i denotes the maximum acceleration of the following vehicle, v_0 denotes the maximum vehicle velocity in free flow traffic, s_0 represents the minimum spacing between two stationary vehicles, T denotes the desired time headway maintained by the following vehicle, b_i denotes the comfortable braking deceleration of the following vehicle, and the exponent δ is usually assumed to have a value of 4 [16].

B. System Dynamics

The self-organized vehicle cluster formation is replicated as a system of M vehicles on a closed ring-road of length L in this work. The closed ring-road environment simplifies the problem by avoiding open boundary conditions due to vehicles entering or exiting the system. For the purposes of this paper, the system under consideration is restricted to $M = 3$ vehicles on a closed ring road of length $L = 60$ m. This simple construction allows for easier visualization of the system dynamics as compared to a large-scale MAS, while still retaining the ability to demonstrate the proposed framework for identifying locally influential agents. For this system, our interest lies in the evolution of the system state described by the state vector $\boldsymbol{x} = [x_1, \dot{x}_1, x_2, \dot{x}_2, x_3, \dot{x}_3]^T \in \mathbb{R}^6$.

The *local controllability* of the system can be determined by linearizing the system about a current operating point ($\boldsymbol{x_0}$), and studying the rank of the controllability matrix. It is assumed that each vehicle has the ability to change its acceleration-related parameters, i.e. a_i (the maximum acceleration) and b_i (the comfortable braking deceleration). By controlling such acceleration-related parameters, the agents can potentially influence the self-organized vehicle cluster formation as seen in [2]. Now, the linearization of the 3-vehicle system yields the local evolution equations given by (4). In this equation, $\alpha_j^{(i)}$ represents the partial derivative of the acceleration of the i^{th} vehicle with respect to the state variable j at the operating point $\boldsymbol{x_0}$, where the acceleration is given by the car-following model in (3). For example, $\alpha_{x_i}^{(1)}$ represents $\partial f / \partial x_i$ evaluated for the first vehicle at the operating point $\boldsymbol{x_0}$, $\alpha_{\dot{x}_{i-1}}^{(2)}$ represents $\partial f / \partial \dot{x}_{i-1}$ evaluated for the second vehicle at the operating point $\boldsymbol{x_0}$, and so on. Similarly, $\beta_k^{(i)}$ represents the partial derivative of the acceleration of the i^{th} vehicle with respect to the input k, evaluated at the operating point $\boldsymbol{x_0}$. For example, $\beta_a^{(1)}$ represents $\partial f / \partial a$ evaluated for the first vehicle at the operating point $\boldsymbol{x_0}$.

$$\underbrace{\begin{bmatrix} \delta\dot{x}_1 \\ \delta\ddot{x}_1 \\ \delta\dot{x}_2 \\ \delta\ddot{x}_2 \\ \delta\dot{x}_3 \\ \delta\ddot{x}_3 \end{bmatrix}}_{\dot{\boldsymbol{x}}} = \underbrace{\left[\begin{array}{cc:cc:cc} 0 & 1 & 0 & 0 & 0 & 0 \\ \alpha_{x_i}^{(1)} & \alpha_{\dot{x}_i}^{(1)} & 0 & 0 & \alpha_{x_{i-1}}^{(1)} & \alpha_{\dot{x}_{i-1}}^{(1)} \\ \hdashline 0 & 0 & 0 & 1 & 0 & 0 \\ \alpha_{x_{i-1}}^{(2)} & \alpha_{\dot{x}_{i-1}}^{(2)} & \alpha_{x_i}^{(2)} & \alpha_{\dot{x}_i}^{(2)} & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \alpha_{x_{i-1}}^{(3)} & \alpha_{\dot{x}_{i-1}}^{(3)} & \alpha_{x_i}^{(3)} & \alpha_{\dot{x}_i}^{(3)} \end{array}\right]}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} \delta x_1 \\ \delta\dot{x}_1 \\ \delta x_2 \\ \delta\dot{x}_2 \\ \delta x_3 \\ \delta\dot{x}_3 \end{bmatrix}}_{\boldsymbol{x}} + \underbrace{\left[\begin{array}{cc:cc:cc} 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_a^{(1)} & \beta_b^{(1)} & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_a^{(2)} & \beta_b^{(2)} & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta_a^{(3)} & \beta_b^{(3)} \end{array}\right]}_{\boldsymbol{B}} \underbrace{\begin{bmatrix} \delta a_1 \\ \delta b_1 \\ \delta a_2 \\ \delta b_2 \\ \delta a_3 \\ \delta b_3 \end{bmatrix}}_{\boldsymbol{u}} \tag{4}$$

C. Studying Agent Influence

Agent influence can be determined by studying the minimum singular value of the augmented matrix $[sI-A, B](s \in \mathbb{C})$, as proposed by Eising in [12]. To study the influence of a specific agent on the system dynamics, the B matrix can be modified appropriately, so that only the specific agent of interest can provide control inputs to the system. For example, in order to measure the influence of vehicle 1 on the system dynamics, one can set $\beta_a^{(2)} = \beta_b^{(2)} = \beta_a^{(3)} = \beta_b^{(3)} = 0$, so that the B matrix reads as:

$$B = B_1 = \left[\begin{array}{cc:cc:cc} 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_a^{(1)} & \beta_b^{(1)} & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{5}$$

In this scenario, only the first agent (or vehicle) may influence the system dynamics via its control actions (or acceleration and deceleration maneuvers). Now, the minimum singular value σ_1 of the augmented matrix $[sI - A, B_1]$ yields a measure of influence of agent 1 on the full-order system dynamics. Comparing the minimum singular value σ_1 with the minimum singular values σ_2 and σ_3 obtained from the corresponding augmented matrices $[sI-A, B_2]$ and $[sI - A, B_3]$ respectively, one can identify the locally most influential agent in the system.

However, the methodology presented up to this point is applicable to all systems, and does not leverage the fact that several multi-agent system exhibit self-organized behavior. Before modifying the presented methodology, we first study the self-organizing behavior of the prototypical system, i.e. self-organized vehicle cluster formation on the closed ring-road. The 3-vehicle system was simulated with model parameters values typically used in the IDM model to model traffic behavior, such as $a = 1$ m/s^2, $b = 3.4$ m/s^2, $T = 2.5$ s, $s_0 = 2$ m, and $v_0 = 15$ m/s. The maximum free flow velocity takes into account the sharp turning radius of the ring road. The results of the simulation are included in Figure 1, which offers offers some evidence that spatio-temporal patterns may be present in the system as it evolves over time. This hypothesis is further strengthened when the phase portrait of the system is observed. However, since the system evolves in $\mathbb{R}^6$, we can only visualize the phase portrait via projections onto lower-dimensional spaces. For such a visualization, the phase portrait is projected into two $\mathbb{R}^3$ spaces, one given by the spacing of the vehicles in the system, i.e. $[s_1, s_2, s_3]^T = [x_1-x_2, x_2-x_3, x_3-x_1]^T \in \mathbb{R}^3$, and the other given by the velocities of the vehicles in the system, i.e. $[v_1, v_2, v_3]^T \in \mathbb{R}^3$. These projections are shown in Figures 2 and 3 and they exhibit behavior similar to a limit cycle, which indicates the presence of a low-dimensional manifold on which the self-organized dynamics of the system evolve. Since the goal of this study is to identify the agents that have most significant influence on the self-organized dynamics, the next step involves determination of the minimum embedding dimension of the low-dimensional manifold on which these dynamics reside.

IV. Minimum Embedding Dimension of Low-Dimensional Manifold

The evidence presented in Figures 2 and 3 indicates that the self-organized dynamics of the multi-agent system may evolve on a low-dimensional manifold in $\mathbb{R}^M$, where $M \ll N$. Knowledge of the dimension of this manifold can assist with the model order reduction of the linearized system equations presented in (4), resulting in a more accurate estimation of the locally influential agents. The use of a reduced-order model also has significant computational advantages for extremely large-scale multi-agent system whose self-organized dynamics may evolve on a manifold of very small dimension.

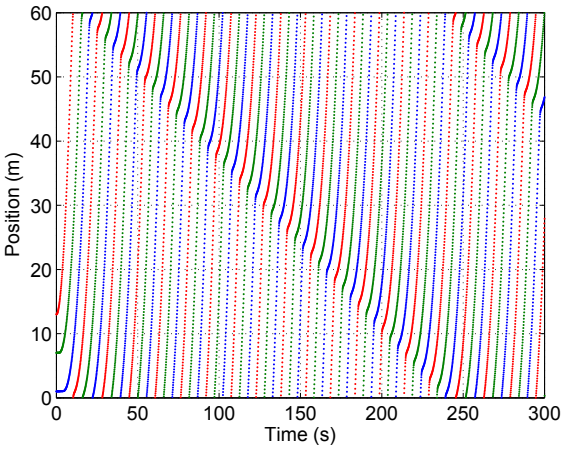

Fig. 1. Three vehicle system on a ring road with dynamics manifesting as self-organized stop-and-go waves.

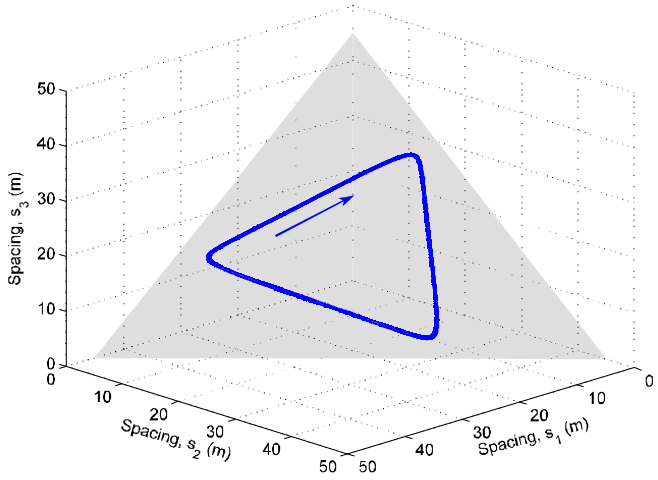

Fig. 2. Behavior similar to a limit-cycle observed in the $\mathbb{R}^3$ projection given by vehicle spacings $[s_1, s_2, s_3]^T$. Arrow indicates the direction of system evolution. The shaded region indicates the plane in which the limit cycle resides.

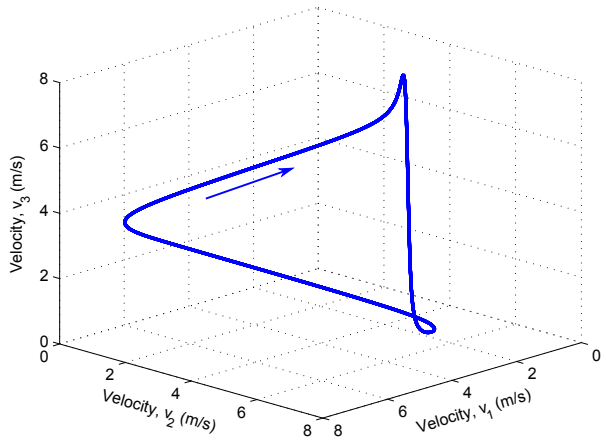

Fig. 3. Behavior similar to a limit-cycle observed in the $\mathbb{R}^3$ projection given by vehicle velocities $[v_1, v_2, v_3]^T$. Arrow indicates the direction of system evolution.

The dimension of the attractor or low-dimensional manifold can be determined by several methods. While the Hausdorff dimension, or its more practical counterpart, the box dimension, may yield accurate estimates of the attractor dimension, these methods require significant computational resources [17]. Moreover, since the ultimate goal is to determine the dimension for a reduced-order model, an upper bound of the attractor dimension serves our purpose equally as well as the exact attractor dimension. Fortunately, there exist methods to determine the upper bound of the attractor dimension that are computationally inexpensive. One such technique is the method of false neighbors developed by Kennel et al. [18]. A modification of this method by Cao is utilized in this paper and is explained in the following subsection.

A. Method of False Neighbors

The underlying premise of the method of false neighbors is that the projection of a system trajectory onto a lower-dimensional space may cause distant points on the trajectory to appear closer than they really are. Such points, which appear to be geometrically-proximate in a lower dimension, may not truly be neighbors in the dynamical sense of the evolution of the system trajectory, and hence are referred to as *false neighbors* [18]. As the system trajectory is projected onto increasingly higher dimensions (though still less than the original dimension of the state space), fewer points on the trajectory are forced into proximity due to geometric constraints, i.e. the number of false neighbors decreases. The dimension at which the percentage of false neighbors drops to zero indicates the minimum embedding dimension d_0, i.e. the attractor of a self-organizing system can be embedded in a state space whose dimension is *at least* d_0. Thus, the minimum embedding dimension can serve as a guide for the reduced-order modeling of large-scale multi-agent systems.

Evaluation of the minimum embedding dimension using Cao's method requires scalar time series data of the state trajectory $\boldsymbol{x}(t) = [x_1, \dot{x}_1, x_2, \dot{x}_2, x_3, \dot{x}_3]^T$. The scalar time series data can be generated from the state trajectory as follows:

$$\chi(t) = c_1 x_1 + c_2 \dot{x}_1 + c_3 x_2 + c_4 \dot{x}_2 + c_5 x_3 + c_6 \dot{x}_3 \qquad (6)$$

where $c_i \in \mathbb{R} \setminus \{0\}$. Without loss of generality, we may assume that $c_i = 1 \; \forall \; i \in \{1, 2, ..., 6\}$, since, in principle, this choice does not affect the evaluation of the minimum embedding dimension. Using Takens' state space reconstruction technique [19], one can generate state trajectories of increasingly higher dimensions until the percentage of false neighbors falls to zero. Specifically, the reconstructed state space is developed in the form of multivariate vectors in d-dimensional space as follows:

$$y_s(d) = [\chi(s), \chi(s+\tau), ..., \chi(s+(d-1)\tau)]^T \qquad (7)$$
$$(s = 1, 2, ..., S - (d-1)\tau)$$

where $y_s(d)$ represents a d-dimensional reconstructed vector generated at time s, τ is the time delay used in the reconstruction (whose choice depends on the system dynamics [20]), and S represents the length of the scalar time series data set. The choice of time delay τ relies on the system dynamics, and is independent of the minimum embedding dimension. For example, using $d = 2$, one can generate a two-dimensional time series $y(s) = [\chi(s), \chi(s+\tau)]^T$. If the dimension of the attractor or low-dimensional manifold is $d_0 = 2$, then all reconstructed vectors of dimension $d > 2$ will ideally have no false neighbors.

However, the method of false neighbors, as originally proposed by Kennel et al., has a few limitations in the form of heuristic choices of thresholds pertaining to what constitutes a 'close' neighbor. Consequently, a modification of the method of false neighbors, as proposed by Cao, is used in the following work [21].

B. Minimum Embedding Dimension via Cao's Method

Cao presented two metrics that draw inspiration from the method of false neighbors, while simultaneously avoiding its limitations. Specifically, a measure of the change in distance between two neighboring points, when the state space is reconstructed using a higher embedding dimension, is evaluated as follows:

$$a(s,d) = \frac{||y_s(d+1) - y_n(d+1)||}{||y_s(d) - y_n(d)||} \tag{8}$$

where $||y_s(d) - y_n(d)||$ represents the ℓ_1 norm used to measures the distance between the d-dimensional reconstructed vector y_s and its neighbor y_n at time s, and $||y_s(d+1) - y_n(d+1)||$ represents the distance between the same points when reconstructed using a $(d+1)$-dimensional space. The value of $a(s,d)$ is large for a specific set of false neighbors that are geometrically close in dimension d, but not in dimension $d+1$. The mean value of $a(s,d)$ over the entire time series data set is evaluated as:

$$E(d) = \frac{1}{S - d\tau} \sum_{s=1}^{S-d\tau} a(s,d) \tag{9}$$

Realizing that various $a(s,d)$ ceases to vary for $d \geq d_0$, where d_0 represents the minimum embedding dimension, it is observed that $E(d)$ also does not vary for dimensions beyond d_0. Cao defines the E_1 metric is defined as follows:

$$E_1(d) = \frac{E(d+1)}{E(d)} \tag{10}$$

From evaluating the E_1 metric for known attractors of dynamical systems such as the Hénon map, it is observed that the metric tends to unity beyond the minimum embedding dimension d_0. Similarly, the E_2 metric defined by Cao also tends to unity beyond the minimum embedding dimension. Due to space constraints, it is not possible to provide a full description of the E_2 metric here, but additional details on this metric can be found in [21].

The minimum embedding dimension is evaluated for the self-organized dynamics of the 3-vehicle system on a closed ring-road. Figure 4 indicates the variation in the E_1 and E_2 metrics for reconstructed vectors $y_s(d)$ of increasingly higher dimensions. It is found that both metrics approach unity at $d = 2$, which implies that the self-organized dynamics of the 3-vehicle system can be completely embedded in a low-dimensional manifold of dimension $d_0 = 2$. The next section discussed how this information can be used in the model order reduction of the linearized system (4).

V. Identification of Influential Agents

As discussed earlier, the goal of this study is to identify the agents that have the ability to influence the self-organized dynamics of the multi-agent system. In Section III, we discussed a scheme for evaluating agent influence for the full-order linearized system. Since the self-organized dynamics evolve on a low-dimensional manifold, the logical next step is to reduce the full-order linearized system and identify the agent influence in the reduced-order model. The minimum embedding dimension determined in the previous section indicates that the state vector of the reduced-order model should have a dimension $d_0 = 2$.

A. Model Order Reduction via Krylov Subspaces

Over the years, several methods have been proposed for the model order reduction of large-scale linear and nonlinear systems, such as balanced truncation and Krylov subspace-based methods, to name a few [22]. Balanced truncation reduces the model order by retaining only the controllable and observable states of the system. While this approach is beneficial in many scenarios, we do not know *a priori* if the controllable states will generate a reduced-order model whose dynamics evolve on the desired low-dimensional manifold. In other words, the relationship between self-organization and controllability is currently unknown and remains an open problem for the research community.

On the other hand, the analytical structure provided by Krylov subspace-based methods can leverage some topological information to obtain a local approximation of the low-dimensional manifold and a reduced-order linearized model. A d-dimensional Krylov subspace for a system such as (4) is defined as follows:

$$\mathcal{K}_d(A, b_0) = \text{span}\{b_0, Ab_0, ..., A^{d-1}b_0\} \tag{11}$$

where $A \in \mathbb{R}^{N\times N}$ represents the system matrix of the multi-agent system linearized about the current operating point $\boldsymbol{x_0}$, and $b_0 \in \mathbb{R}^{N\times 1}$ represents a *starting vector* used to generate the Krylov subspace [23][24]. The choice of the starting vector is critical in generating a reduced-order model that accurately depicts the system dynamics. Ideally, the starting vector should lie in the invariant subspace in which the dynamics evolve, so that the Krylov subspace can capture the system behavior. Typically, this information is not available, but the current context presents insights which can guide the selection of the starting vector. Specifically, the tangent

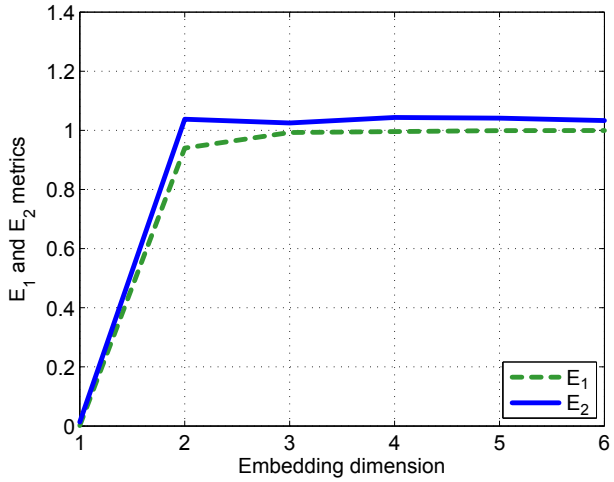

Fig. 4. Minimum embedding dimension for the self-organized dynamics of the 3-vehicle system is found to be $d_0 = 2$. The minimum embedding dimension is determined via Cao's method as the dimension at which E_1 and E_2 approach unity.

to the system trajectory is expected to lie in the local low-dimensional manifold of the system equations. This tangent vector, defined as:

$$b_0 = \left[\frac{\partial f^{(1)}}{\partial x_n}, \frac{\partial f^{(1)}}{\partial \dot{x}_n}, \frac{\partial f^{(2)}}{\partial x_n}, \frac{\partial f^{(2)}}{\partial \dot{x}_n}, \frac{\partial f^{(3)}}{\partial x_n}, \frac{\partial f^{(3)}}{\partial \dot{x}_n}\right]^T \quad (12)$$

can be evaluated at each time instant t at the current operating point $\boldsymbol{x_0}$ on the system trajectory. The tangent vector can then serve as the starting vector of the Krylov subspace for model order reduction at each time instant as we traverse over the system trajectory.

An orthonormal basis of the Krylov subspace $\mathcal{K}_d(A, b_0)$ may be used as a projection operator $U = [u_1, u_2, ..., u_d]$ $(u_i \in \mathbb{R}^N)$ to project the dynamics of the full-order system described in (4) onto a local low-dimensional manifold. In the current context, since the minimum embedding dimension is $d_0 = 2$, the associated projection operator $U \in \mathbb{R}^{6\times 2}$ obtained by Arnoldi's process is given by:

$$U = \left[\frac{b_0}{||b_0||}, \frac{b_1}{||b_1||}\right] \quad (13)$$

$$\text{with } b_1 = Ab_0 - \frac{\langle b_0, Ab_0\rangle}{\langle b_0, b_0\rangle} b_0 \quad (14)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. The next subsection discusses how this information may be used to identify agents that can influence the self-organized dynamics of the 3-vehicle system.

B. Influential Agents in 3-Vehicle System

The reduced-order model of the system described in (4) may be obtained through the operations $A_r = U^T A U$ and $B_r = U^T B_i$, so that the reduced-order linearized system equations are:

$$\dot{\xi} = A_r \xi + B_r u \quad (15)$$

where $\boldsymbol{x} = U\xi$, and $B_i \in \{B_1, B_2, B_3\}$ depending on the agent i whose influence on the self-organizing dynamics is being assessed. The minimum singular value is evaluated at the current operating point $\boldsymbol{x_0}$ over the system trajectory for the three different augmented matrices $[sI - A_r, U^T B_1]$, $[sI - A_r, U^T B_2]$, and $[sI - A_r, U^T B_3]$, corresponding to control inputs provided only by vehicle 1, vehicle 2, or vehicle 3, respectively. Figure 5 describes the state trajectories as well as the trends in the minimum singular value for each system configuration. The linearized reduced-order system that has the largest minimum singular value helps identify the agent that exerts the most influence on the self-organized dynamics at a particular instant of time.

Analysis of Figure 5 yields some interesting insights that agree strongly with intuition about the 3-vehicle system, and about macroscopic traffic jams in general. For example, it is observed that vehicles with large spacings, which represent vehicles in free flow, do not have significant influence on the self-organized dynamics. Additionally, it is also observed that vehicles that are about to enter a cluster (i.e. vehicles whose spacing is decreasing with time), have significant influence

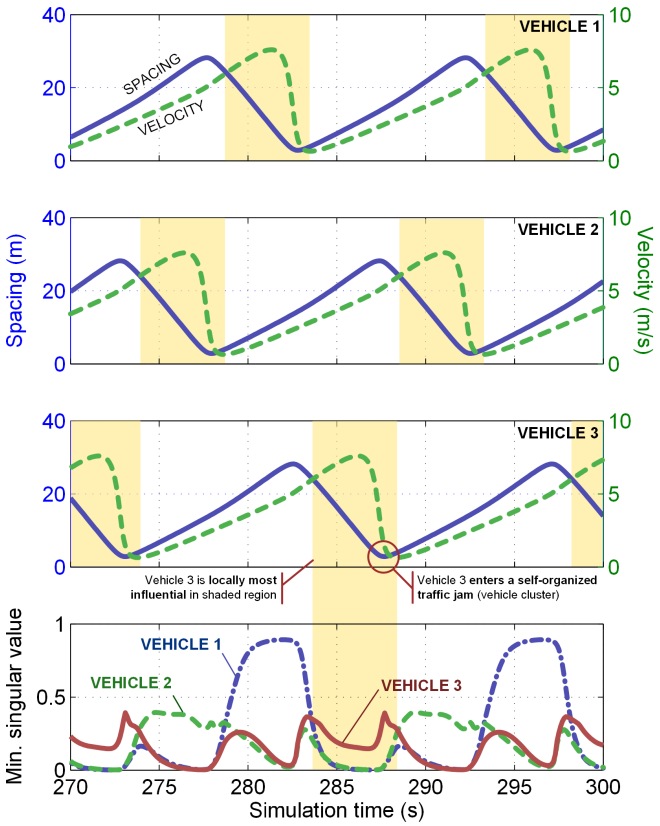

Fig. 5. State trajectories for the 3-vehicle system. For configurations where different vehicles are designated as controllers, the largest minimum singular values help identify influential agents. Shaded yellow regions indicate the spatial locations where a particular agent (or vehicle) has the most local influence on the self-organized dynamics of the system.

over the self-organized dynamics. These results indicate a strong spatial dependence of agent influence and agree with intuition as well as similar results from previous related research on connected vehicles [25].

VI. Conclusions

This paper proposed a methodology for assessing the influence of specific agents on the self-organized dynamics of a large-scale multi-agent system. The methodology made use of existing knowledge of controllability metrics and model order reduction, and coupled it with the novel application of the minimum embedding dimension to identify agent influence on self-organized dynamics. Approaching the problem from a topological perspective also provided a solution to the selection of an appropriate starting vector for the Krylov subspace-based model order reduction process.

The methodology presented here is general enough to be applied to any self-organizing system. The technique was applied to a known prototypical problem of self-organizing multi-agent systems, i.e. the self-organized vehicle cluster formation on closed ring roads. The obtained results which identify the most influential agents at any instance of time were found to agree with intuition as well as related previous studies. Specifically, it was found that vehicles in free flow

have little influence on the self-organized dynamics, whereas vehicles approaching a cluster are quite influential in this regard. The generality of the proposed approach hopes to propel further research directed towards assessing the role of agent influence in guiding self-organizing behavior.

REFERENCES

[1] H. Haken, *Information and self-organization*. Springer-Verlag, 1988.

[2] K. Jerath and S. N. Brennan, "Analytical prediction of self-organized traffic jams as a function of increasing ACC penetration," *IEEE Trans. on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1782–1791, 2012.

[3] D. Boison, "Adenosine kinase, epilepsy and stroke: mechanisms and therapies." *Trends in pharmacological sciences*, vol. 27, no. 12, pp. 652–658, Dec 2006.

[4] S. Sinha and K. A. Siddiqui, "Definition of intractable epilepsy." *Neurosciences*, vol. 16, no. 1, pp. 3–9, Jan 2011.

[5] X. F. Wang and G. Chen, "Pinning control of scale-free dynamical networks," *Physica A*, vol. 310, pp. 521–531, 2002.

[6] M. Porfiri and M. Di Bernardo, "Criteria for global pinning-controllability of complex networks," *Automatica*, vol. 44, no. 12, pp. 3100–3106, 2008.

[7] S. Patterson and B. Bamieh, "Leader selection for optimal network coherence," in *Proceedings of the 49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 2692–2697.

[8] H. Kawashima and M. Egerstedt, "Leader selection via the manipulability of leader-follower networks," in *American Control Conference (ACC), 2012*. IEEE, 2012, pp. 6053–6058.

[9] A. Clark, L. Bushnell, and R. Poovendran, "On leader selection for performance and controllability in multi-agent systems." in *Proceedings of the 51st IEEE Conference on Decision and Control (CDC)*, 2012, pp. 86–93.

[10] A. Chapman and M. Mesbahi, "System theoretic aspects of influenced consensus: Single input case," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1505–1511, June 2012.

[11] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, March 2014.

[12] R. Eising, "Between controllable and uncontrollable," *Systems & Control Letters*, vol. 4, no. 5, pp. 263–264, Jul 1984.

[13] A. Hamdan and A. Elabdalla, "Geometric measures of modal controllability and observability of power system models," *Electric Power Systems Research*, vol. 15, no. 2, pp. 147–155, 1988.

[14] C. N. Viswanathan, R. Longman, and P. Likins, "A degree of controllability definition - fundamental concepts and application to modal systems," *Journal of Guidance, Control, and Dynamics*, vol. 7, no. 2, pp. 222–230, Mar 1984.

[15] K. B. Lim, "Method for optimal actuator and sensor placement for large flexible structures," *Journal of Guidance, Control, and Dynamics*, vol. 15, no. 1, pp. 49–57, Jan 1992.

[16] A. Kesting, M. Treiber, and D. Helbing, "Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1928, pp. 4585–4605, 2010.

[17] C. Beck and F. Schögl, *Thermodynamics of Chaotic Systems: An Introduction*. Cambridge University Press, 1995.

[18] M. B. Kennel, R. Brown, and H. D. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.

[19] F. Takens, *Dynamical systems and turbulence, Warwick 1980*. Springer, 1981, vol. 898, ch. Detecting strange attractors in turbulence, pp. 366–381.

[20] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical review A*, vol. 33, no. 2, pp. 1134–1140, 1986.

[21] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, no. 1–2, pp. 43–50, Dec 1997.

[22] A. Antoulas, D. Sorensen, and S. Gugercin, "A survey of model reduction methods for large-scale systems," *Contemporary mathematics*, vol. 280, pp. 193–220, 2001.

[23] B. Lohmann and B. Salimbahrami, "Introduction to krylov subspace methods in model order reduction," *Methods and Applications in Automation*, pp. 1–13, 2000.

[24] D. S. Watkins, *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods*. Society for Industrial and Applied Mathematics, 2007, ch. Krylov Subspace Methods.

[25] K. Jerath, V. V. Gayah, and S. N. Brennan, "Influential subspaces of connected vehicles in highway traffic," in *Proceedings of Symposium Celebrating 50 Years of Traffic Flow Theory, Portland, OR*, 2014.