Modes of Information Flow in Collective Cohesion

Sulimon Sattari
Research Center of Mathematics for Social Creativity,
Research Institute for Electronic Science, Hokkaido University,
Kita 20, Nishi 10, Kita-ku, Sapporo 001-0020, Japan

Udoy S. Basak
Graduate School of Life Science, Transdisciplinary Life Science Course,
Hokkaido University, Kita 12, Nishi 6, Kita-ku, Sapporo 060-0812, Japan and
Pabna University of Science and Technology, Pabna 6600, Bangladesh

James P. Crutchfield
Complexity Sciences Center and Physics Department,
University of California at Davis, One Shields Avenue, Davis, California 95616, USA

Tamiki Komatsuzaki
Research Center of Mathematics for Social Creativity,
Research Institute for Electronic Science, Hokkaido University,
Kita 20, Nishi 10, Kita-ku, Sapporo 001-0020, Japan
Institute for Chemical Reaction Design and Discovery (WPI-ICReDD),
Hokkaido University Kita 21 Nishi 10, Kita-ku, Sapporo, Hokkaido 001-0021, Japan and
Graduate School of Chemical Sciences and Engineering Materials Chemistry and Engineering Course,
Hokkaido University, Kita 13, Nishi 8, Kita-ku, Sapporo 060-0812, Japan
(Dated: December 2, 2020)

Pairwise interactions between individuals are taken as fundamental drivers of collective behavior—responsible for group cohesion and decision-making. While an individual directly influences only a few neighbors, over time indirect influences penetrate a much larger group. The abiding question is how this spread of influence comes to affect the collective. In this, one or a few individuals are often identified as "leaders", being more influential than others in determining group behaviors. To support these observations transfer entropy and time-delayed mutual information are used to quantitatively identify underlying asymmetric interactions, such as leader-follower classification in aggregated individuals—cells, birds, fish, and animals. However, these informational measures do not properly characterize asymmetric interactions. They also conflate distinct functional modes of information flow between individuals and between individuals and the collective. Employing simple models of interacting self-propelled particles, we examine the pitfalls of using them to quantify the strength of influence from a leader to a follower. Surprisingly, one must be wary of these pitfalls even for two interacting particles. As an alternative we decompose transfer entropy and time-delayed mutual information into intrinsic, shared, and synergistic modes of information flow. The result not only properly reveals the underlying effective interactions, but also facilitates a more detailed diagnosis of how individual interactions lead to collective behavior. This exposes, for example, the role of individual and group memory in collective behaviors.

I. INTRODUCTION

Coherent collective behavior fascinates us when a global pattern emerges from individuals who share information with others only in their local vicinity. Decisions made by one individual apparently cascade throughout the entire group. That said, not all members have the same influence. The challenge here in explaining such emergent behaviors is to infer the underlying relationships among individuals from observations. Not surprisingly, this challenge has attracted many over decades to diagnose collective behaviors in a variety of systems [4, 20, 23, 25, 26, 29, 31, 37]. In epithelial Madin-Darby canine kidney monolayers, for example, collective cell migration is triggered by multicellular protrusions forming a fingerlike structure [25, 26, 37]. Photoablating a group of cells from the finger tip makes the remaining cell group lose its sense of direction. The interpretation is that the former and latter cells have acted as if they were "leaders" and "followers". This functional assignment can be carried out due to their spatial location along the protrusion and since the two cell classes are genetically distinct [37].

Identifying leaders—even defining what that role means [9]—is very difficult. Especially so, when probing the mechanisms that cause the dynamical behaviors of aggregated agents. Beyond cells, these basic questions also apply to bird flocking [20], fish schooling [4, 23], caribou migration [31], and bonobo ape foraging [29]. Leader agents are often defined as an individual that exerts more influence upon others than others influence on upon it. That is, the role is fundamentally asymmetric. Previous studies [1, 4, 17–19, 23] proposed that, under this definition, pairwise analysis of trajectories can assign leaders and followers under the working hypothesis that a change in motion of the leader forecasts a change in motion of the follower. From this, one interprets the change of leader motion as a candidate cause that triggers the motion of followers.

Various statistical quantities are used to infer causal relationships in such situations. In pigeon flocks, for example, the time-delayed correlation between the orientation of individuals at one time instance and the orientation of others at previous times reveals a hierarchical leadership structure and also provides a method to quantify the timescale of influence [20]. In such a case, the motion of one pigeon is correlated with the past motion of another. Granger causality [3] is seen as an improvement to time-delayed correlation as it quantifies the predictability of the current state of a variable based on knowledge of a variable at a previous time. Time-delayed correlation and Granger causality both assume linear relationships between variables, though. This generally does not hold. More recent studies argued that information-theoretic quantities—transfer entropy, time-delayed mutual information, and causation entropy—are superior when identifying leaders since they naturally accommodate the highly nonlinear nature of multi-agent systems [1, 4, 13, 14, 17, 19, 22, 30].

In principle, trajectories of each agent are required to fully analyze the mutual influence of individuals at desired spatiotemporal resolution. However, in most of empirical settings one cannot make such detailed observations. Here, we assume that only pairwise measurements—that is, coordinates of two agents including their history—are available to quantify the influence of individuals in collective motion. Carefully considering the definitions of (formally pairwise) information-theoretic quantities—such as, transfer entropy or time-delayed mutual information—further illuminates the types of influence a particular individual has. As pointed out in introducing transfer entropy [27], time-delayed mutual information reports a nonzero value between the present of a stochastic variable X and the future of a stochastic variable Y even when Y has no direct influence on X. This implies that it cannot be directly employed to infer the underlying mutual influence among individuals. It also includes additional information not intrinsically coming from X.

Transfer entropy, in contrast, computes the reduction of uncertainty about Y's future while knowing X's present, conditioned on Y's present. Recently though, Ref. [11] showed that, paralleling time-delayed mutual information, transfer entropy incorporates additional, unwarranted information; namely, the reduction of uncertainty about Y that occurs by knowing the present state of X and Y simultaneously. This information is extraneous to determining "flow" and, misleadingly, adds to the desired information: intrinsic flow from X to Y. In this view, transfer entropy decomposes into two distinct modes of information flow—intrinsic and synergistic [11].

Pairwise interactions are fundamental to information theory's development of input-output ("two-port") communication channels [7]. As such, they provide a primary statistical tool that, as we show, makes it possible to infer the underlying influences among individuals. To obtain maximum insight into the mechanisms underlying multi-agent systems, the following focuses on decomposing transfer entropy and time-delayed mutual information into Ref. [11]'s three different fundamental modes of information flow—termed intrinsic, shared, and synergistic information flows. The results demonstrate how the decomposed elemental information flows shed light on the influences that drive leader-follower relationships.

As an illustrative vehicle we employ a generalized Vicsek model [33] with two additional features: 1) tunable influence weight of one particle over another (i.e., leaders have larger influence) and 2) particle memory. We show that, by analyzing the effects of 1) and 2) on the three modes of information flow, intrinsic information flow exists whenever the motion of an agent depends on another with nonzero weight, as does transfer entropy and time-delayed mutual information. Shared and synergistic information, however, can occur when agents mutually influence each other and synergistic information only occurs in such cases. These results extend previous studies on modes of information flow in finite-state hidden Markov models by introducing Vicsek models that are fundamental to understanding collective motion and exhibit distinct modes of information flow. Moreover, we probe the effect of agent memory and the effect of more than two interacting agents on the different modes of information flow and their role in collective behavior.

II. RESULTS

A. Inferring causal relationship: Time-delayed mutual information and transfer entropy

In our definition of leaders, leaders are expected to be more influential than followers. Due to the influence exerted by the leader on a follower, the follower's motion at some time is influenced by the leader's, but there exists a disparity in time. The field of information theory has introduced widely used measures such as time-delayed mutual information and transfer entropy to characterize causal relationships between variables. Such quantities go beyond time-delayed correlation [17] or Granger causality [3], which only capture relationships at the linear level. Consider that $X = (..., x_{t-1}, x_t, x_{t+1}, ...)$ and $Y = (..., y_{t-1}, y_t, y_{t+1}, ...)$ are two stationary processes with probability mass functions $p(x_t) = Pr\{X = x_t\}$ and $p(y_t) = Pr\{Y = y_t\}$, respectively. Time-delayed mutual information (TDMI), denoted by $\mathcal{M}_{X\to Y}$ hereinafter, is defined as [13]:

$$\mathcal{M}_{X\to Y}(\tau) = I(X_t; Y_{t+\tau}) = H(Y_{t+\tau}) - H(Y_{t+\tau}|X_t) = H(X_t) - H(X_t|Y_{t+\tau}), \tag{1}$$

where, for example, $H(Y_{t+\tau})$ and $H(Y_{t+\tau}|X_t)$ are the Shannon entropy and the conditional entropy, respectively, which estimate the uncertainty of the variable $Y_{t+\tau}$ and the remaining uncertainty of the variable $Y_{t+\tau}$ once given X_t with a delay time τ. In other words, the mutual information is the reduction of uncertainty in the future of Y by knowing the present of X at a present time t. Since mutual information is symmetric, this can also be interpreted as the reduction of uncertainty in X at a present time t by knowing Y in the future time $t+\tau$, however, this interpretation cannot be used to infer causality since future cannot influence present.

As Schreiber elucidated [27], TDMI has a drawback in predicting influence, because the quantity can be nonzero when two variables have some shared history. That is, the condition $I(X_t; Y_{t+\tau}) > 0$ may hold even when variable Y is not directly influenced by variable X when either X or Y dynamics contains memory of their configurations.

To overcome the above-mentioned shortcomings, transfer entropy (TE) has been introduced [27]. The idea of TE is that if a process X influences another process Y, then the prediction of future Y becomes easier after knowing the present of both X and Y, compared to only knowing the present of Y. TE from X to Y, denoted by $\mathcal{T}_{X\to Y}$ hereinafter, has the following form:

$$\mathcal{T}_{X\to Y}(\tau) = I(X_t; Y_{t+\tau}|Y_t) = H(Y_{t+\tau}|Y_t) - H(Y_{t+\tau}|Y_t, X_t), \tag{2}$$

that is, $\mathcal{T}_{X\to Y}(\tau)$ is time-delayed mutual information between Y at time $t+\tau$ and X at time t conditioned by Y at time t and is exactly the same as the subtraction of the remained uncertainty in Y at time $t+\tau$ given both X and Y at the present time t from that in $Y_{t+\tau}$ given Y_t, which corresponds to the amount of uncertainty of $Y_{t+\tau}$ reduced by knowing X_t in addition to the knowledge of Y_t. TE has been used extensively as one of the most standard methods for measuring influence such as in classifying leaders and followers [5, 14, 17–19, 22, 23], and to infer causal relationships systems pervading many areas of science including neuroscience [28, 32, 34–36], chemistry [2], human behavior [21, 24], and Earth systems [6, 8, 10], as it is an improvement upon TDMI for quantifying asymmetric causal relationships between variables. It is noted, however, that, like correlation, information-theoretic quantities such as TDMI and TE are not enough in themselves to identify causality. The potential existence of hidden variables should be taken into consideration. Since the conditioning on the past in computing TE and TDMI are both finite in time length, the history of each variable may act like another hidden variable which influences the outcome, leading to a spurious effect in the computation of information flow, as we will elucidate further.

Recently, it was demonstrated for a simple toy binary system that positive TE from X to Y can exist even though knowledge of X_t alone cannot reduce the uncertainty of $Y_{t+\tau}$ in the system [11] (See also SI). It was pointed out that, in addition to information intrinsic to the reduction of uncertainty in the variable $Y_{t+\tau}$ by knowing the present of X_t being independent of the present of Y_t, TE from X to Y includes the information corresponding to the reduction of uncertainty in the variable $Y_{t+\tau}$ by knowing the present of X_t and Y_t simultaneously [12]. Here the idea of information flow being intrinsic is defined as the amount of uncertainty that knowledge of the present of X alone reduces about the future of Y further than what knowing the present of Y alone or knowing the present of X and Y simultaneously would reduce. As a potential measure to avoid computing influence that is coming from both the present of X and Y in predicting the future of Y, intrinsic mutual information has been proposed [12, 16]. Note that while intrinsic mutual information is a specific quantity and is not synonymous with intrinsic information flow, intrinsic mutual information is an attempt to compute intrinsic information flow between two variables.

B. Inferring causal relationship: Intrinsic, shared, and synergistic mutual information

In order to explain the motivation behind intrinsic mutual information, one must take the cryptographic flow ansatz [12], which states that intrinsic information flow between X and Y is synonymous with secret key agreement [15, 16] between X_t and $Y_{t+\tau}$ while Y_t is an outside observer. Based on this ansatz, information which is flowing intrinsically from X to Y is equal to the secret key agreement rate $S(X_t; Y_{t+\tau}|Y_t)$. Since $S(X_t; Y_{t+\tau}|Y_t)$ is not computable in practice, intrinsic mutual information $\mathcal{I}_{X\to Y}$ is used as a convenient upper bound on $S(X_t; Y_{t+\tau}|Y_t)$ to demonstrate information which is flowing intrinsically from X to Y. Intrinsic mutual information from a process $X = \{..., x_{t-1}, x_t, x_{t+1}, ...\}$ to another process $Y = \{..., y_{t-1}, y_t, y_{t+1}, ...\}$, denoted by $\mathcal{I}_{X\to Y}$, is defined as the infimum of $I(X_t; Y_{t+\tau}|\overline{Y}_t)$, taken over all possible conditional distributions $p(\overline{Y}_t|Y_t)$ [11] (See also SI), i.e.,

$$\mathcal{I}_{X\to Y}(\tau) := \inf\Big\{I(X_t; Y_{t+\tau}|\overline{Y}_t) : \sum_{y\in Y_t} p(X_t, Y_{t+\tau}, Y_t = y)p(\overline{Y_t}|Y_t = y)\Big\}, \tag{3}$$

where $\overline{Y_t}$ is an auxiliary variable to elucidate the upper bound of $S(X_t; Y_{t+\tau}|Y_t)$ that satisfies a Markov property $X_tY_{t+\tau} \to Y_t \to \overline{Y_t}$. Here $A \to B$ ($AC \to B$) signifies that B depends only on A (A and/or C) and the infimum is taken over all possible conditional distributions $p(\overline{Y_t}|Y_t)$. Intrinsic mutual information $\mathcal{I}_{X\to Y}$ represents uncertainty reduction in the future of Y by knowing only the present of X as much as possible under the assumption of the Markov property with respect to Y and $\overline{Y}$. It should be noted here that this Markov property has no physical relevance to the actual system Y, and is merely an assumption to reduce the space of the minimization in Eq. 3, and provide an upper bound of $S(X_t; Y_{t+\tau}|Y_t)$. Intrinsic mutual information has been found to be both a convenient and accurate bound on the secret key agreement rate $S(X_t; Y_{t+\tau}|Y_t)$ [12, 16]. The following relations sum up the importance of $\mathcal{I}_{X\to Y}(\tau)$ and its relationship to $S(X_t; Y_{t+\tau}|Y_t)$, $\mathcal{T}_{X\to Y}(\tau)$, and $M_{X\to Y}(\tau)$:

$$0 \le S(X_t; Y_{t+\tau}|Y_t) \le \mathcal{I}_{X\to Y}(\tau), \tag{4}$$

$$\mathcal{I}_{X\to Y} \le \mathcal{T}_{X\to Y}(\tau), \tag{5}$$

and

$$\mathcal{I}_{X\to Y} \le \mathcal{M}_{X\to Y}(\tau). \tag{6}$$

Equations 4-6 demonstrate the utility of intrinsic mutual information that it can be used to compute bounds on the deviations of $\mathcal{T}_{X\to Y}(\tau)$ and $\mathcal{T}_{X\to Y}(\tau)$ from $S(X_t; Y_{t+\tau}|Y_t)$. That is, whenever equality does not hold in Eq. 5 (Eq. 6), then there must be a portion of $\mathcal{T}_{X\to Y}(\tau)$ ($\mathcal{M}_{X\to Y}(\tau)$) which is not intrinsically coming from X. From here onwards we set $\tau = 1$ and omit τ from the equations, as it has been shown that $\tau = 1$ best captures the information flow between two particles in the Vicsek model [1].

Once $\mathcal{I}_{X\to Y}$ has been computed, the subtractions from TDMI ($\mathcal{M}_{X\to Y}$) and TE ($\mathcal{T}_{X\to Y}$) are defined and denoted by $\sigma_{X\to Y}$ and $\mathcal{S}_{X\to Y}$, respectively:

$$\sigma_{X\to Y} = \mathcal{M}_{X\to Y} - \mathcal{I}_{X\to Y}, \tag{7}$$

and

$$\mathcal{S}_{X\to Y} = \mathcal{T}_{X\to Y} - \mathcal{I}_{X\to Y}. \tag{8}$$

Based on Eqs. 4-6, one can deduce that $0 \le \sigma_{X\to Y} \le \mathcal{M}_{X\to Y}$ and $0 \le \mathcal{S}_{X\to Y} \le \mathcal{T}_{X\to Y}$. Since $\mathcal{I}_{X\to Y}$ serves as the information coming from (mostly) X alone (since intrinsic mutual information provides an upper bound of $S(X_t; Y_{t+\tau}|Y_t)$) to Y, $\mathcal{S}_{X\to Y} > 0$ ($\sigma_{X\to Y} > 0$) implies that $\mathcal{T}_{X\to Y}$ ($\mathcal{M}_{X\to Y}$) contains some information that comes from the present of Y, and $\mathcal{T}_{X\to Y}$ ($\mathcal{M}_{X\to Y}$) should not be treated as information flowing only from X to Y in those cases. Here, based on the interpretation that $\mathcal{I}_{X\to Y}$ is (mostly) the information coming from only X to Y, $\sigma_{X\to Y}$ is the part of $\mathcal{M}_{X\to Y}$ which is coming from knowledge of both variables, and is postulated to be the information redundantly shared by both X and Y, thus, termed as “shared” information. Likewise $\mathcal{S}_{X\to Y}$ is the part of $\mathcal{T}_{X\to Y}$ which is coming from knowledge of both variables, and is postulated to be the information coming from simultaneous knowledge of X and Y, termed as “synergistic” information [11]. These three quantities $\mathcal{I}$, σ, and $\mathcal{S}$ carry a more detailed picture of the relationship between X and Y than can be inferred by computing $\mathcal{T}$ or $\mathcal{M}$ alone. In this paper, we show using modified Vicsek model that $\mathcal{T}$ and $\mathcal{M}$ can result in misleading interpretation concerning the underlying actual relationship among individuals, and propose that $\mathcal{I}$, σ, and $\mathcal{S}$ can provide us with a more firm interpretation of the relationship without requiring additional experiments.

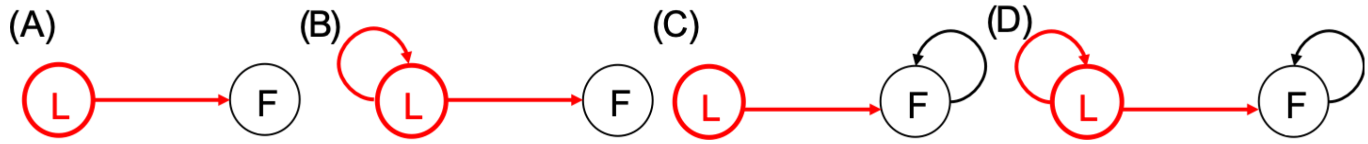

FIG. 1. Graph representation of four example interaction types with asymmetric interaction between leader and follower, i.e., $w_{\mathrm{LF}} > 0$ and $w_{\mathrm{FL}} = 0$. a) The dynamics of the leader and follower do not depend on their present configurations. b) Only the dynamics of the leader depends on its present configuration. c) Only the dynamics of the follower depends on its present configuration. d) The dynamics of the leader and the follower both depend on their own present configuration. Note that the value of w_{LL} (w_{FF}) is inconsequential when the dynamics of leader (follower) do not depend on its present configuration, and the $\theta_i(t)$ depends solely on a random number in the interval $[-\pi, \pi]$.

C. Modified Vicsek Model

To demonstrate the drawback of $\mathcal{T}$ or $\mathcal{M}$, and the interpretability of their decompositions into modes $\mathcal{I}$, σ, and $\mathcal{S}$, we employ a series of modified Vicsek models that modify the original one [33] with asymmetric interaction weights and turn on/off dependence on the present dynamics of interacting particle(s) in the propagation of system dynamics. Consider N particles lying within a square box of length L with periodic boundary conditions. The position of each particle i ($i = 1, 2, ..., N$) at time t is denoted by $\vec{r}_i^{\,t}$ and is updated with a time increment Δt as:

$$\vec{r}_i^{\,t+1} = \vec{r}_i^{\,t} + \vec{v}_i^{\,t} \Delta t, \tag{9}$$

where $\vec{v}_i^{\,t}$ denotes the velocity of particle i at time t. For simplicity, we consider particles with a uniform constant speed v_0, and only their orientations θ_i change. The orientation of a particle is updated at each time step by taking the weighted average of the velocity of neighboring particles within a given radius R:

$$\theta_i(t+1) = \langle \boldsymbol{\theta}(t) \rangle_{R, \boldsymbol{w}, \vec{\boldsymbol{r}}^t} + \Delta\theta_i, \tag{10}$$

where $\langle \boldsymbol{\theta}(t) \rangle_{R, \boldsymbol{w}, \vec{\boldsymbol{r}}^t}$ is computed by $\arctan\left[\sum_j' w_{ij} \sin\theta_j(t) / \sum_j' w_{ij} \cos\theta_j(t)\right]$ where $\sum'$ takes over all j satisfying $|\ \vec{r}_i^{\,t} - \vec{r}_j^{\,t}\ | \leq R$. $\boldsymbol{w}(= \{w_{ij}\})$ is a nonnegative asymmetric matrix whose element in the i^{th} row and j^{th} column represents the interaction strength that particle i exerts on particle j, and $w_{ij} > w_{ji}$ whenever particle i is a leader and particle j is a follower in our setting. $\Delta\theta_i$ is a random number uniformly distributed in the range $[-\eta_0/2, \eta_0/2]$ and is chosen uniquely for each particle i at each time step and represents thermal noise. In the original Vicsek model, the right hand side of Eq.(10) ensures that the dynamics of $\theta_i(t+1)$ result from the configurations of all the particles j (including that of the same particle i) within the circle of radius R centered at $\vec{r}_i^{\,t}$.

We construct a set of the modified Vicsek models by turning on and off the dependence on $\theta_j(t)$ associated with follower/leader in determining the $\theta_i(t+1)$ dynamics, where the leader influences the follower but the follower does not influence the leader (i.e., $w_{\mathrm{LF}} > 0$ while $w_{\mathrm{FL}} = 0$). Figure 1 shows graph representations of the influence of particles in a simple, two particle system in determining $\theta_i(t+1)$ in four different cases, where L and F denote leader and follower, respectively, and A $\rightarrow$ B (A, B is either L or F) signifies that the present of A influences the future of B. Here, we vary $w_{\mathrm{LF}} (\in \mathbb{N})$ from 1 to 10, and set w_{LL} and w_{FF}=1 to be unity for the models in which the present of L(F) influences the future of L(F). For models where no influence of the present to the future of the same particle (F and/or L) exists (Figs. 1A- 1C), we replaced $\theta_i(t)$ in the summation $\sum'$ by a random number in the interval $[-\pi, \pi]$, to erase any influence from the present of $\theta_i(t)$ of the same particle i.

Misinterpretation of causal inference in time-delayed mutual information (TDMI) and transfer entropy (TE) for a two-particle system: Let us first examine the TE ($\mathcal{T}$) and TDMI ($\mathcal{M}$) landscapes shown in Figures 2 and 3 as a function of w_{LF} and η_0 for different interaction types in Fig. 1, where $\mathcal{T}$ and $\mathcal{M}$ from leader to follower (L$\rightarrow$F) are given in A-D and those from follower to leader (F$\rightarrow$L) are in E-H. As expected, $\mathcal{M}_{\mathrm{L}\rightarrow\mathrm{F}}$ (Fig. 2A-2D) and $\mathcal{T}_{\mathrm{L}\rightarrow\mathrm{F}}$ (Fig. 3A- 3D) exhibit significant, nonzero values in all interaction types, since the leader is influencing the follower in all cases. Naively one should expect $\mathcal{M}_{\mathrm{F}\rightarrow\mathrm{L}}$ and $\mathcal{T}_{\mathrm{F}\rightarrow\mathrm{L}}$ to be zero for all cases since the follower does not influence the leader at all, however, there are in fact spurious values of $\mathcal{M}_{\mathrm{F}\rightarrow\mathrm{L}}$ (Figs 2F and 2H) and $\mathcal{T}_{\mathrm{F}\rightarrow\mathrm{L}}$ (Figs. 3F and 3H) whenever the future dynamics of leader depends on its present. This is due to the leader's history acting as a hidden variable that imparts information onto both the leader and the follower. In the following paragraph and in Sects II D- II F, we elaborate on the effect of the leader and follower dynamics that are dependent on their present on the various information flow quantities ($\mathcal{T}$, $\mathcal{M}$, and their decompositions into $\mathcal{I}$, σ, and $\mathcal{S}$).

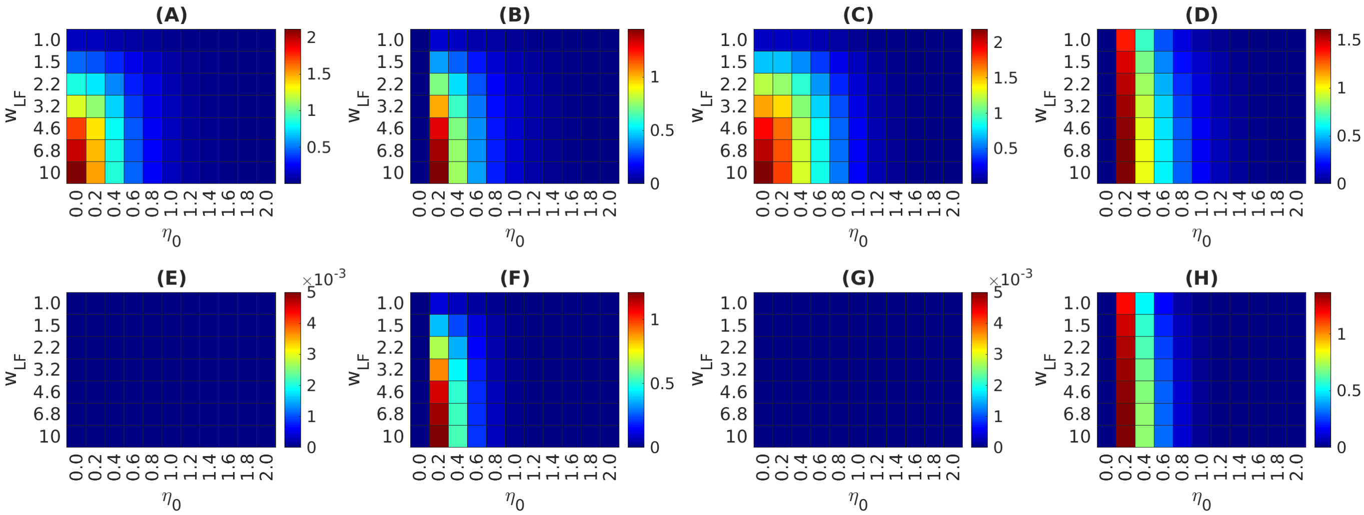

FIG. 2. Time-delayed mutual information (TDMI) $\mathcal{M}_{\mathrm{L}\to\mathrm{F}}$ (A-D) and $\mathcal{M}_{\mathrm{F}\to\mathrm{L}}$ (E-H) as a function of the strength of interaction w_{LF} and the noise level η_0. A) and E) for type A, B) and F) for type B, C) and G) for type C, and D) and H) for type D as diagrammed in Fig. 1.

Decomposed information modes, intrinsic, shared, and synergistic information: Figures 4, 5, and 6 show different modes of information, i.e., intrinsic ($\mathcal{I}$), shared (σ), and synergistic informations ($\mathcal{S}$) defined by Eqs. (3-8) from leader to follower and vice versa, for four different interaction types given in Fig. 1, respectively. First, let us look into the decomposed modes of information for each interaction types in brief. Each interaction diagram corresponds to a special case in information flow. In type A where neither the leader nor the follower depends on its present in its $\theta_i(t+1)$ dynamics (Fig. 1A), only $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ and no other type of information flow is observed as seen in Figs. 4A, 5A, and 6A. Fig. 1B represents type B where only the leader depends on its present in the $\theta_i(t+1)$ dynamics. This is similar to the example given by Schreiber [27], where, perhaps counterintuitively, there exists a significant amount of TDMI from the follower to the leader $\mathcal{M}_{\mathrm{F}\to\mathrm{L}}$ even though there exists no direct interaction in that direction, as seen in Fig. 2F. Transfer entropy $\mathcal{T}$ was then introduced to reconcile this issue [27], and $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ does in fact reduce the amount information flow in that direction in our model, given that $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ is less than $\mathcal{M}_{\mathrm{F}\to\mathrm{L}}$ in Figs. 3F and 2F (we will elucidate where this spurious TE is coming from in Sect II F). Fig. 1C represents type C where only the follower depends on its present in the $\theta_i(t+1)$ dynamics. This tends to erase all spurious values of information flow from follower to leader (Figs. 2G-6G), since those spurious effects are coming from the history of the leader, as we will elucidate further in the following sections. The leader to follower information flows in this case (Figs. 2C-6C) are all increased compared to the case where the follower dynamics do not depend on its present (Figs. 2A-6A), which further emphasizes our point that dependence of the dynamics on its present plays a key role in the calculation of information flow even when the interactions between individuals are not intrinsically changing (we elucidate further in the following sections). Figure 1D may represent the most intuitive case in typical physical systems, where both entities depend on their present in the $\theta_i(t+1)$ dynamics. In that case, all flows have significant values except for the synergistic information from the follower to the leader $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$, which is negligibly small (Fig. 6H). In the following, we look deeper into each mode of information for their illustrative types of interactions.

D. Intrinsic information

$\mathcal{I}_{X\to Y}$ is the uncertainty reduction in future of Y from knowing the present of (mostly) X alone. Hence $\mathcal{I}$ is more fundamental than TE or TDMI. $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ (Fig. 4A- 4D) is nonzero while the larger the noise level η_0 the lower the value of $\mathcal{I}_{L\to F}$ and $\mathcal{I}_{F\to L}$. $\mathcal{I}_{L\to F}$ increases as w_{LF} increases, as one may expect. In the cases where the dynamics of the leader depends on its present configuration in types B and D (Figs. 4B and 4D), the leader generates less information at $\eta_0 = 0$ because its dynamics are dependent largely on its initial conditions at time $t = 0$ in trajectories, and therefore imparts less information upon the follower. As η_0 increases, the leader experiences a much wider variety of configurations from time to time and, hence, generates more information and therefore $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ increases as a function of η_0 at low η_0. However as η_0 approaches 2π, noise dominates the follower's motion and therefore the follower's motion depends less on the leader's, and $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ goes back to zero. In turn, in the cases where the dynamics of the leader does

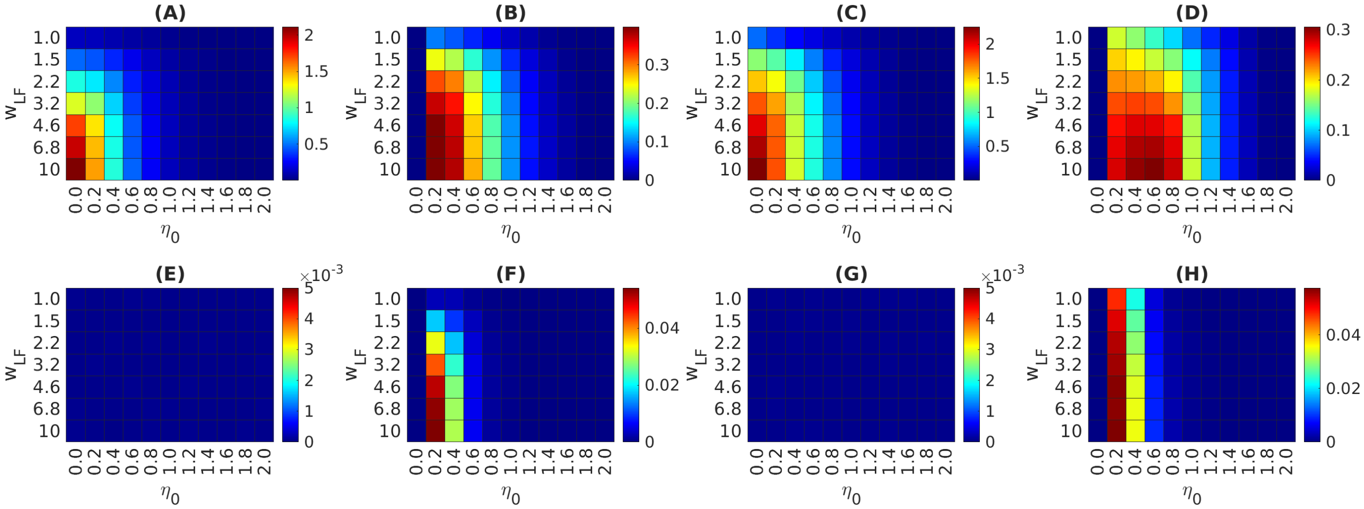

FIG. 3. Transfer entropy (TE) $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ (A-D) and $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ (E-H) as a function of the strength of interaction w_{LF} and the noise level η_0. A) and E) for type A, B) and F) for type B, C) and G) for type C, and D) and H) for type D as diagrammed in Fig. 1.

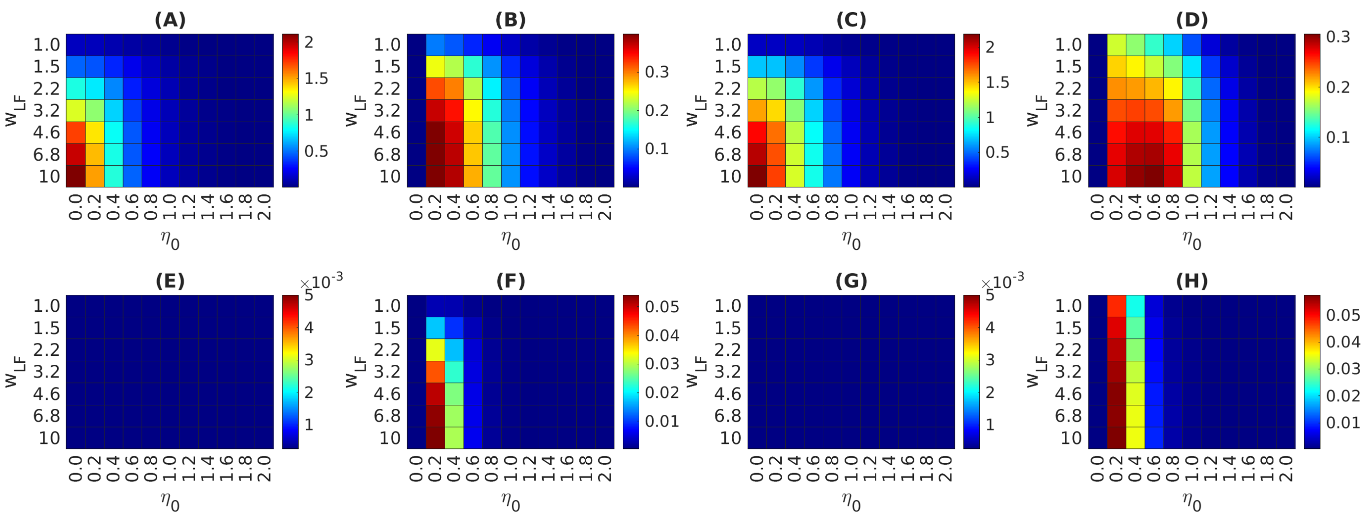

FIG. 4. Intrinsic information $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ (A-D) and $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$ (E-H) as a function of the strength of interaction w_{LF} and the noise level η_0. A) and E) for type A, B) and F) for type B, C) and G) for type C, and D) and H) for type D as diagrammed in Fig. 1.

not depend on its present in types A and C (Figs. 4A and 4C), $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ peaks at $\eta_0 = 0$. This arises from the fact that the simulation procedure erases any influence of $\theta_i(t)$ in determining the future state $\theta_i(t+1)$ irrespective of the value of η_0, which randomizes the leader configuration at each time step after interacting with the follower (hence making the leader impart new information to the follower), and therefore information transfer is highest when the thermal noise on the follower is lowest and its motion is dominated by the interaction.

$\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$ is zero for types A and C where the leader dynamics do not depend on its present as shown in Figs. 4E and 4G. However, note for types B and D where the leader dynamics depend on its present, $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$ spuriously exists even though there is no direct interaction from the follower to the leader (Figs. 4F and 4H). As clarified below in Figs. 6F and 6H, $\mathcal{T}_{\mathrm{F}\to\mathrm{L}} \simeq \mathcal{I}_{\mathrm{F}\to\mathrm{L}}$ and this is the same situation as observed in the TE landscape (Figs. 3F and 3H). Note that the computation of intrinsic information $\mathcal{I}$ (Eq. 3) makes $\mathcal{I}_{X\to Y}$ least depend on the history of Y with an auxiliary variable $\overline{Y}$ to mimic Y along its Markov property: $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$ is regarded as minimizing the possible influence of the leader's history to determining the future state of the leader $\theta_{\mathrm{L}}(t+1)$.

Thus the history of the leader's configuration may act as a hidden variable which injects information to both the follower and the leader. To elucidate this, we have computed types B and D only with the leader forgetting its past once every two time steps, which is truly Markovian in its leader dynamics. That is, the leader's dynamics at time t only depend on its dynamics at time $t-1$ when t is even, otherwise its orientation at time $t-1$ in Eq. 10 is replaced

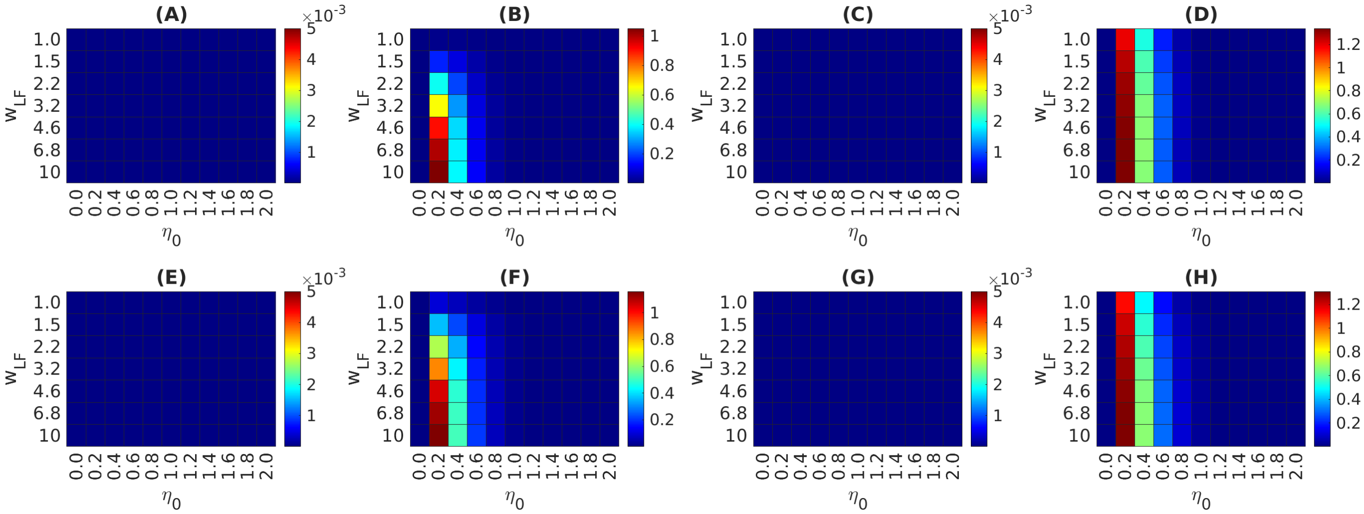

FIG. 5. Shared information $\sigma_{\mathrm{L}\to\mathrm{F}}$ (A-D) and $\sigma_{\mathrm{F}\to\mathrm{L}}$ (E-H) as a function of the strength of interaction w_{LF} and the noise level η_0. A) and E) for type A, B) and F) for type B, C) and G) for type C, and D) and H) for type D as diagrammed in Fig. 1.

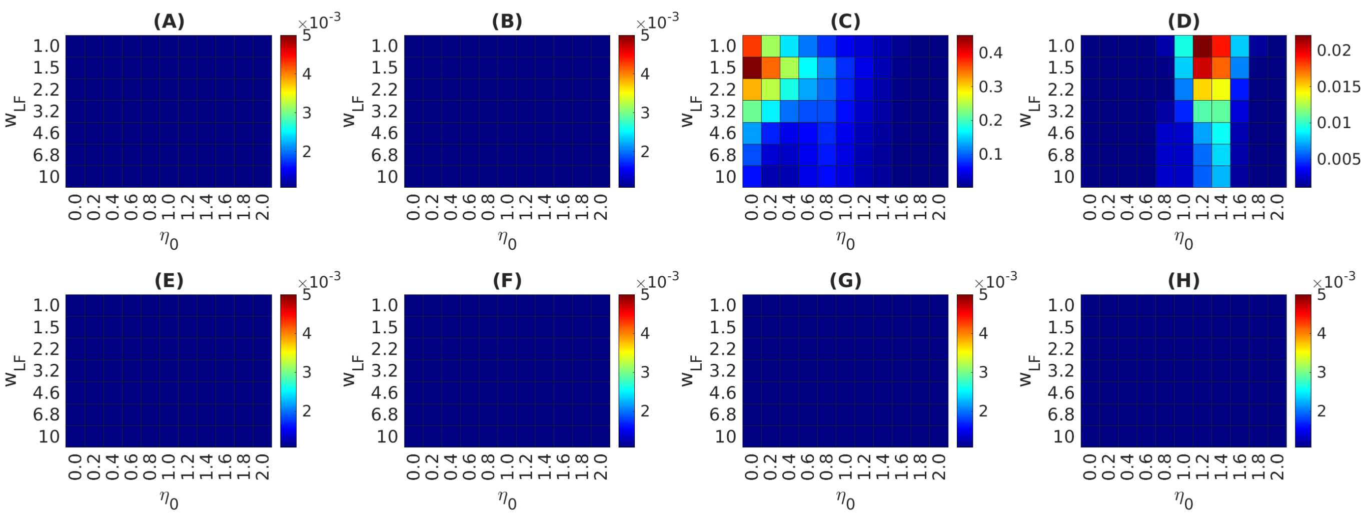

FIG. 6. Synergistic information $\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$ (A-D) and $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$ (E-H) as a function of the strength of interaction w_{LF} and the noise level η_0. A) and E) for type A, B) and F) for type B, C) and G) for type C, and D) and H) for type D as diagrammed in Fig. 1.

by a random number in the interval $[0, 2\pi]$. The reason for this is to remove the possibility that spurious amounts of $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ result from the past of the leader acting as a hidden variable that affects both the leader and the follower. As shown in Figs. 13D and 12D, $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$ becomes negligible in these cases. One can deduce that in the cases where the follower dynamics depends on its past (Figs. 4C and 4D) a small part of $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ is also coming from this effect.

E. Shared information

$\sigma_{X\to Y}$ quantifies the information which is redundantly flowing from both the present of X and Y to the future of Y, and represents the part of $\mathcal{M}_{X\to Y}$ which is not intrinsically coming from X (see Eq. 1). $\sigma_{\mathrm{L}\to\mathrm{F}}$ is negligibly small when the leader does not depend on the present in its dynamics for types A and C, as shown in Figs. 5A and 5C, because the $\sigma_{\mathrm{L}\to\mathrm{F}}$ is coming from information that the leader shares with its present that it has imparted onto the follower. As in the case of $\mathcal{I}$, σ is also zero when $\eta_0 = 0$ since there is no uncertainty to be reduced, as shown in Figs. 5B and 5D. $\sigma_{\mathrm{L}\to\mathrm{F}}$ and $\sigma_{\mathrm{F}\to\mathrm{L}}$ are both high whenever the leader dynamics depends on its present values as more present information is shared in types D and B (Figs. 5B, 5D, 5F, and 5H) and are highest when when both L and F are dependent on their present (Figs. 5D and 5H). In types B and D where σ is nonzero, both $\sigma_{\mathrm{L}\to\mathrm{F}}$ and $\sigma_{\mathrm{F}\to\mathrm{L}}$ have

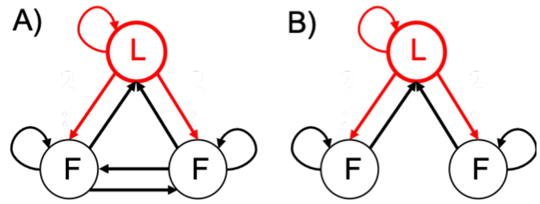

FIG. 7. Interaction diagrams for three agents. A) Model A. Interaction diagram where a leader influences both followers and both followers influence the leader, and followers can also influence each other. B) Model B. Followers cannot influence each other. Weights are asymmetric where w_{LF} (leader to followers) is greater than w_{FL} and w_{FF}. We set w_{LF}=1.05, and $w_{\mathrm{FL}} = w_{\mathrm{FF}} = 1$.

a tendency to increase as a function of w_{LF}, since the future dynamics of the follower depends more on that of the leader, i.e., larger shared information for type D than for type B.

In the case of $\sigma_{\mathrm{F}\to\mathrm{L}}$, it follows a very similar trend as the $\sigma_{\mathrm{L}\to\mathrm{F}}$. In the case where the leader has memory, the values of σ in Fig. 5F imply that TDMI exists between the follower and the leader, as Schreiber had pointed out [27], even though the follower has no direct influence on the leader (in this case, $\mathcal{M}_{\mathrm{F}\to\mathrm{L}} \simeq \sigma_{\mathrm{F}\to\mathrm{L}}$ due to Eq. 1 since $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$ relatively negligible, as seen in Fig. 4F). Thus σ elucidates the part of TDMI which is not intrinsically coming from the leader or from another hidden variable such as shared history.

To elucidate the claim that $\sigma_{\mathrm{F}\to\mathrm{L}}$ comes from the history of the leader, we have computed σ when the leader forgets its memory once every two time steps for types B and D. $\sigma_{\mathrm{F}\to\mathrm{L}}$ (Figs. 13E and 12E) is zero since $\sigma_{\mathrm{F}\to\mathrm{L}}$ conditions on the past of the leader for one time step, and the leader does not have additional memory past one time step.

F. Synergistic information

$\mathcal{S}_{X\to Y}$ represents the information which comes to the future of Y from simultaneously knowing the present of X and the present of Y. When $\mathcal{S}_{X\to Y}$ is high compared to $\mathcal{I}_{X\to Y}$, it means that the future state of Y is only predictable when the present of both X and Y are known simultaneously than when either of them are known independently. $\mathcal{S}_{X\to Y}$ is postulated to be the part of $\mathcal{T}_{X\to Y}$ which arises from the contribution from the present of X and Y simultaneously. (see Eq. 2). Thus one can see that $\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$ is close to zero whenever the follower cannot predict itself, such as when the follower dynamics does not depend on the present of the follower configuration, and is entirely dependent on noise and the leader's dynamics (types A and B), as in Figs. 6A and 6B. When the follower does depend on its present, $\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$ is highest relative to $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$ when the influence of both the leader and follower are balanced (as in Fig. 6D when w_{LF} is close to 1), and therefore knowledge of both leader and follower are required to accurately predict the outcome of the follower. $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$ is zero in all cases except where the leader forgets its past once every two time steps. The nonzero value of $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$ in cases where the leader forgets its past once every two time steps (Figs. 13F and 12F) can be explained as follows. Whenever t is even, it means that the leader has memory of its past at time $t-1$, which it has imparted upon the follower. Whenever $t-1$ is odd, the leader does not have memory of its past and is not likely to be aligned with the follower. Therefore the simultaneous knowledge of $\theta_{\mathrm{L}}(t-1)$ and $\theta_{\mathrm{F}}(t-1)$ informs us about the state of $t-1$ (i.e., $t-1$ is more likely to be even if $\theta_{\mathrm{L}}(t-1)$ and $\theta_{\mathrm{F}}(t-1)$ are aligned). Knowing the status of $t-1$, the follower can then predict whether the leader will correlate with itself at time t or not, and thus can predict the leader's present with some likelihood whenever $t-1$ is even.

G. Leader and group of followers

Up to this point we have considered only pairwise interactions, which is in line with the theoretical basis of the information measures used, for example $\mathcal{T}_{X\to Y}(\tau)$ in Eq. 2. There exist generalizations of these measures to take into account the influence of additional time series of the third particle (or agent) such as causation entropy [14]. However in practice, conditioning on other variables increases the dimension of the probability space required for computing the information measures, and large amounts of data are required in order to properly sample the probability distributions. Therefore, in many-agent systems, it is not usually feasible to condition on all other agents that interact with a given agent. Nonetheless, computation of information measures between two agents has been proven useful for characterizing influence [4, 13, 17, 19, 22], even in systems having more than two agents. We will now show how looking at $\mathcal{I}$, σ, and $\mathcal{S}$, can improve upon the interpretation given the approximate nature of these measures.

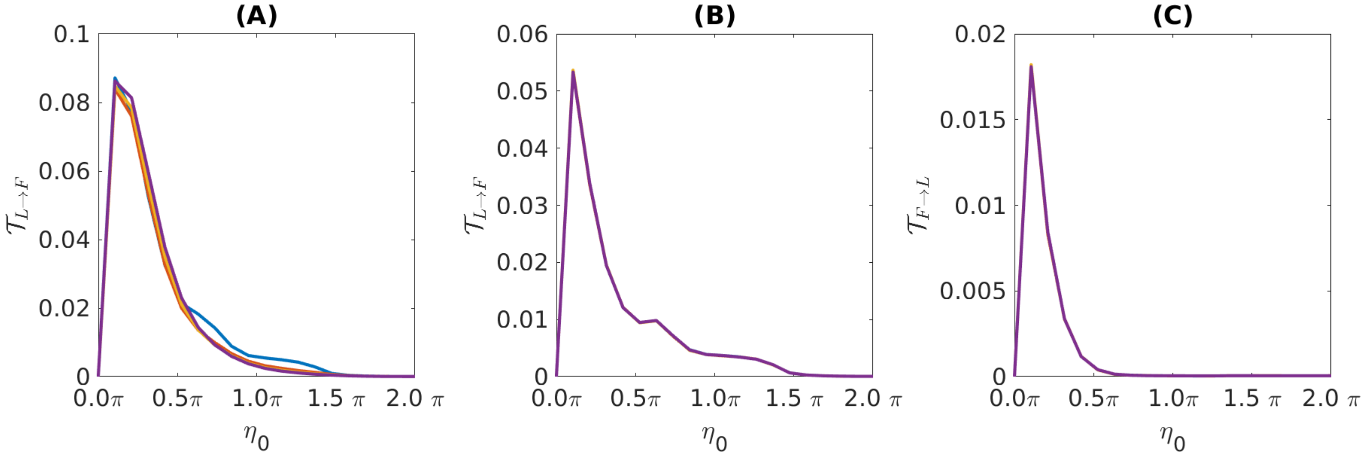

FIG. 8. $\mathcal{T}$ as a function of noise level η_0 with 2 agents (blue), 4 agents (red), 6 agents (orange) and 8 agents (purple) where the number of leaders is always one. A) $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ for model A, B) $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ for model B, and C) $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ for model B.

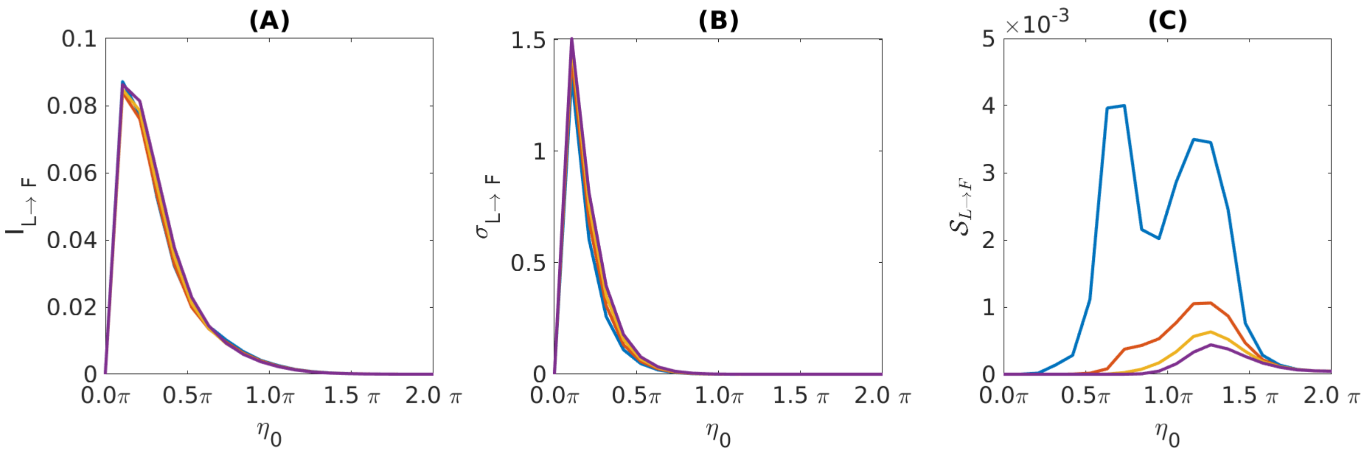

FIG. 9. Modes of information flow from leader to follower the case where followers can influence followers for model A as a function of noise level η_0 with 2 agents (blue), 4 agents (red), 6 agents (orange) and 8 agents (purple). A) $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$, B) $\sigma_{\mathrm{L}\to\mathrm{F}}$, and C) $\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$.

Let us consider the case of the Vicsek model where the leader and follower mutually interact with one another, where model A allows followers to also directly interact with each other but model B does not (e.g., Fig. 7A and Fig. 7B for the case of three agents case). Figures 8A, 8B, and 8C show $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ as a function of noise level for model A, $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ for model B, and $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ for model B, respectively for different numbers of agents in which only one agent is a leader and all the rest are followers (The plot of $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ for model A was almost indistinguishable with Fig. 8A. see Supporting information Fig. S15). As an overall trend, the $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ and $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ first increase shortly and then decrease as the noise level η_0 increases, with small bumps at $0.6\pi \leq \eta_0 \leq 1.5\pi$ for $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ (See Figs. 8A and B). At noise level η_0 being zero, agent's movements quickly attain a regular laminar flow irrespective of what initial configurations and velocities are prepared (see the movie in SI) so that influence of orientational movements of leader or follower is negligible (on average) for predicting orientational motions of the others. In nature, all agents are subject to finite noise due to thermal fluctuation from the environment. Gradual decreases of $\mathcal{T}$ as the noise level gets larger simply arise from the stochastic nature of the system dynamics. However, it should be noted that the gradual decrease of $\mathcal{T}$ for model A (Fig. 8A) is slightly dependent on the number of followers, i.e., third agent(s) for elucidating the pairwise information flow, in which the more the number of followers the (slightly) larger the $\mathcal{T}$ but almost identical irrespective of the number of followers for model B (Figs. 8B and 8C). As seen in Fig. 8A and 8B, the bumps in $\mathcal{T}_{\mathrm{L}\to\mathrm{F}}$ appear in $0.6\pi \leq \eta_0 \leq 1.5\pi$ (as well as in $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ (Fig.15)) for model A start to cease as the number of followers increase while the bumps observed in a similar η_0 region for model B stay unchanged with respect to the increase of the number of followers: in that the smallest number of followers, i.e., just one follower, gives rise to the largest bumps for model A.

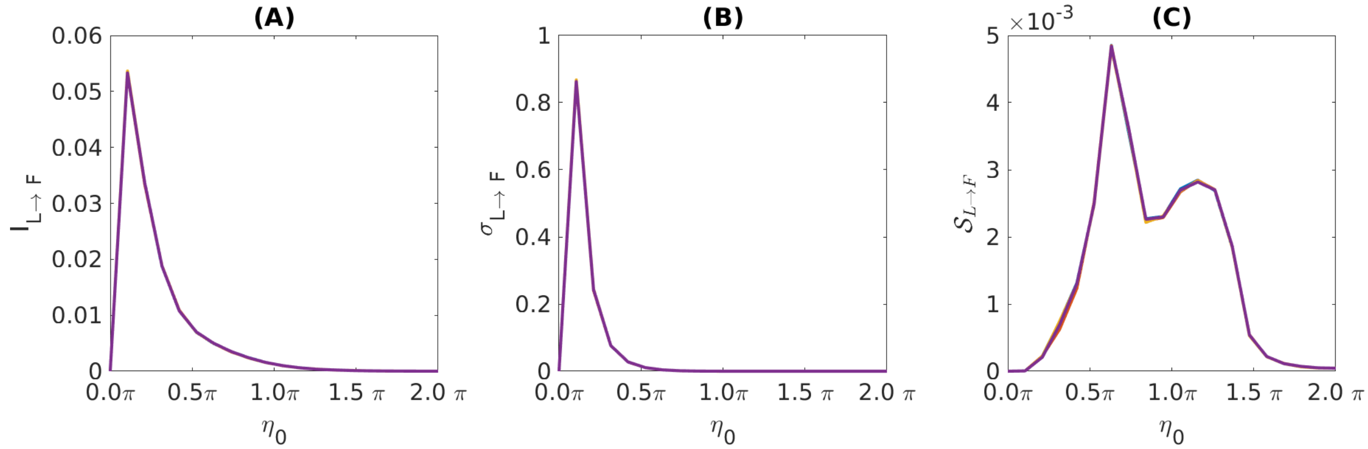

FIG. 10. Modes of information flow from leader to follower for the case where followers cannot influence followers for model B as a function of noise level η_0 with 2 agents (blue), 4 agents (red), 6 agents (orange) and 8 agents (purple). A) $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$, B) $\sigma_{\mathrm{L}\to\mathrm{F}}$, and C) $\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$

The three different modes of 'pairwise' information flow ($\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$, $\sigma_{\mathrm{L}\to\mathrm{F}}$, and $\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$) for model A, and for model B are shown in Fig. 9 and Fig. 10, respectively. One can see that the bumps observed in the plot of $\mathcal{T}$ vs η_0 originate from the ingredient inside transfer entropy, synergistic information $\mathcal{S}$, and the major contribution of transfer entropy $\mathcal{T}_{X\to Y}$ is that of $\mathcal{I}_{X\to Y}$ that is maximally free from the present of Y in the VM model systems.

$\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$ and $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$ for model A (Fig. 9C and Fig. 14C) decrease as the number of followers increases. This is due to the fact that $\mathcal{S} > 0$ only when the simultaneous knowledge of two agents brings additional predictability which is not found in knowing either of the agents alone. As the dynamics depends on more agents, the knowledge of the present state of two agents becomes less powerful in predicting the future state of either agent because its future state is more influenced by the third agent(s). Figure 10 shows the case where we turn off follower-follower interactions. In this case, there is no effect of increasing the number of agents on $\mathcal{I}$ or σ (Figs. 10A, and B and Figs. 14 A and B)). $\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$ (Fig. 10C) also has no effect as the number of agents increases, since, from the follower's point of view, it is only interacting with the leader and is not affected by any other agents. $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$ (Fig. 14 C), however, still decreases as a function of the number of agents, since each follower can influence the leader and therefore increasing the number of followers decreases the likelihood that simultaneously knowing the state of a particular follower and the leader has any additional predictive power on the leader.

It is noted that to confirm the significance of the values of modes of information flow, especially that of $\mathcal{S}$, we computed $\mathcal{I}$, σ, and $\mathcal{S}$ between two stochastic, fully random variables taken from $[0 : 2\pi]$ with the same time length of the leader-follower VM simulation: $\mathcal{I} = 9.2 \times 10^{-6} \pm 1.8 \times 10^{-6}$, $\sigma = 6.0 \times 10^{-17} \pm 7.9 \times 10^{-17}$, and $\mathcal{S} = 4.7 \times 10^{-5} \pm 6.8 \times 10^{-6}$. That is, the above discussions on $\mathcal{S}$, e.g., dependence on the number of followers, persists and not buried in possible number fluctuation in the computation.

III. CONCLUSION

We investigated a series of model systems based on the Vicsek Model of collective motion to explore the effect of interaction protocols on the distinct modes of information flow. In theory, one would condition on all variables, as well as history, to fully interpret mutual relationships among agents in a collective. At present this is not practical. Instead, our task was to acquire detailed and correct interpretations under the constraint of limited measurements—specifically, pairwise interactions among agents.

We observed that the intrinsic information between X and Y dominates whenever there is only a link from X to Y and no direct link between Y to X or from Y to itself. However, a small amount of intrinsic information can still be observed when there is no direct link from X to Y, as in the case where X is a follower with memory and Y is a leader. We noted that this was due to the effect of memory. We also found that shared information dominates when there is memory shared between particles due to their interactions and memory of their pasts. Synergistic information dominates when present knowledge of X or Y alone cannot predict the future state of Y by itself, but knowing both the present of X and Y simultaneously does. One of the most striking consequences in our analysis of this multi-agent system was that decomposing transfer entropy into intrinsic information and synergistic information flows enabled us to correctly interpret the "bump" observed in transfer entropy as a function of noise level. Notably, from that one

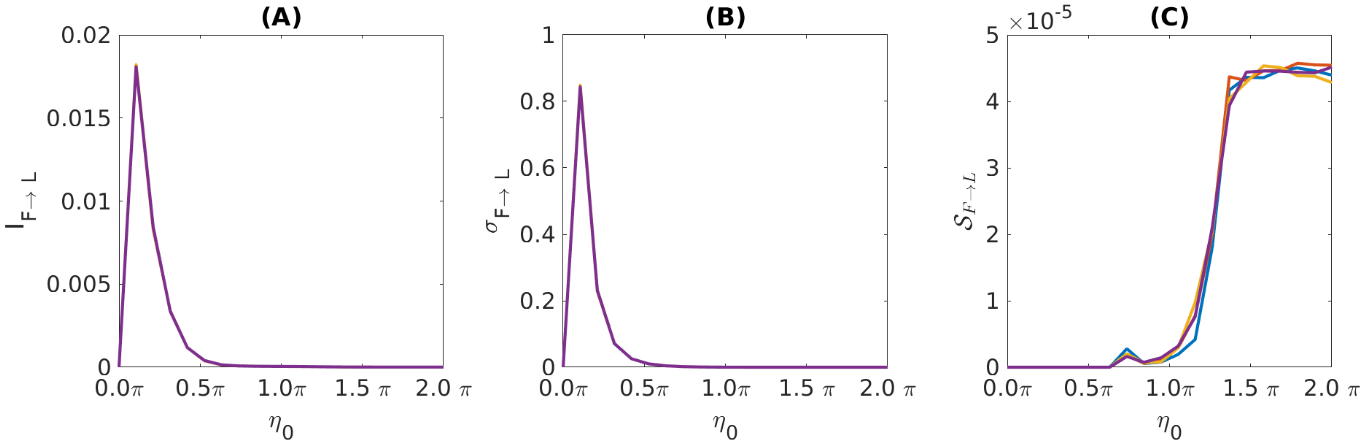

FIG. 11. Modes of information flow from follower to leader for the case where followers cannot influence followers for model B as a function of noise level η_0 with 2 agents (blue), 4 agents (red), 6 agents (orange) and 8 agents (purple). A) $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$, B) $\sigma_{\mathrm{F}\to\mathrm{L}}$, and C) $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$.

can infer how each follower interacts with each other in the collective.

Based on the model systems and their corresponding information flows, one can deduce which information measure is more appropriate based on the physical problem being addressed. In leader-follower classification, for example, transfer entropy is often used. However, when one does not expect to find significant synergistic and shared flows, it equals the time-delayed mutual information. The latter is then a better choice since it does not require additional conditioning that increases the dimension of the probability distribution that must be well-sampled. In cases where synergistic flow is dominant, one may consider separating intrinsic and synergistic flows instead of computing just transfer entropy. This results in a much richer feature space for classification. In general, computing intrinsic, shared, and synergistic flows should perform better or at least as well as transfer entropy and time-delayed mutual information in classification. Future work will verify these claims and elucidate exactly in which scenarios we expect each mode of information flow to be effective in classifying leaders and followers.

In empirical settings, it would be useful to know when there is sufficient time-series data to compute information flows to adequate accuracy, accounting for experimental errors, variables that are not conditioned upon, and history length that is not conditioned upon. Using general formulas for intrinsic information, transfer entropy, and time-delayed mutual information, one can compute the modes of information flow as a function of history length and the number of conditioning variables. In the hope of making the methods more applicable to experiments in a rigorous sense, another future direction for this work is quantifying exactly, and providing bounds for, the amount of intrinsic information that arises from external variables or from shared history.

ACKNOWLEDGMENTS

We thank Profs. Hiroshi Teramoto, J. Green, M. Toda, and Dr. Ryan James for valuable discussions. TK and JPC thank the Telluride Science Research Center for its hospitality during visits. This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas "Singularity Biology (No.8007)" (18H05413), MEXT, and by JSPS (No. 25287105 and 25650044 to T.K.), and JST/CREST (No. JPMJCR1662 to T.K.). It is also based upon work supported by, or in part by, the U. S. Army Research Laboratory and the U. S. Army Research Office under grant W911NF-18-1-0028.

IV. METHODS

A. Details of the modified Vicsek Model

Positions are updated using Eq. 9, and the orientations θ_{F} and θ_{L} are updated using Eq. 10. The weighted interaction term in Eq. 10 is computed by

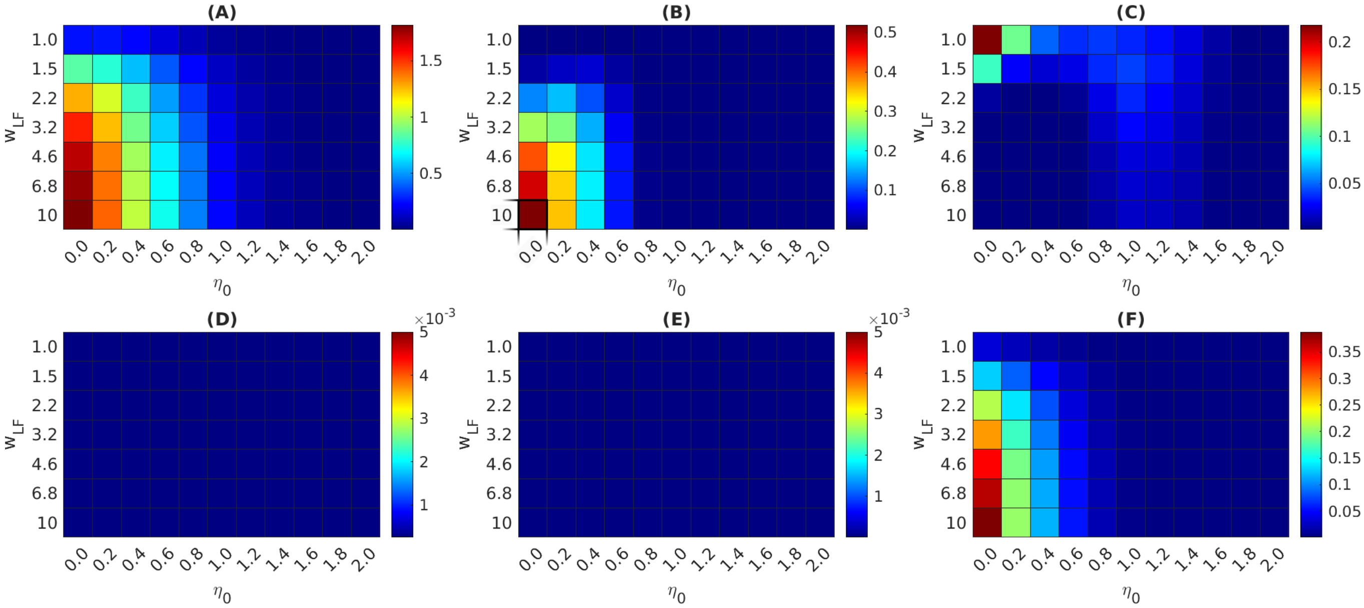

FIG. 12. (A) $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$, (B)$\sigma_{\mathrm{L}\to\mathrm{F}}$, (C)$\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$, (D) $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$, (E)$\sigma_{\mathrm{F}\to\mathrm{L}}$, and (F)$\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$ for interaction type D in Fig. 1, only with the leader forgetting its past once every two time steps. (we should move this to SI)

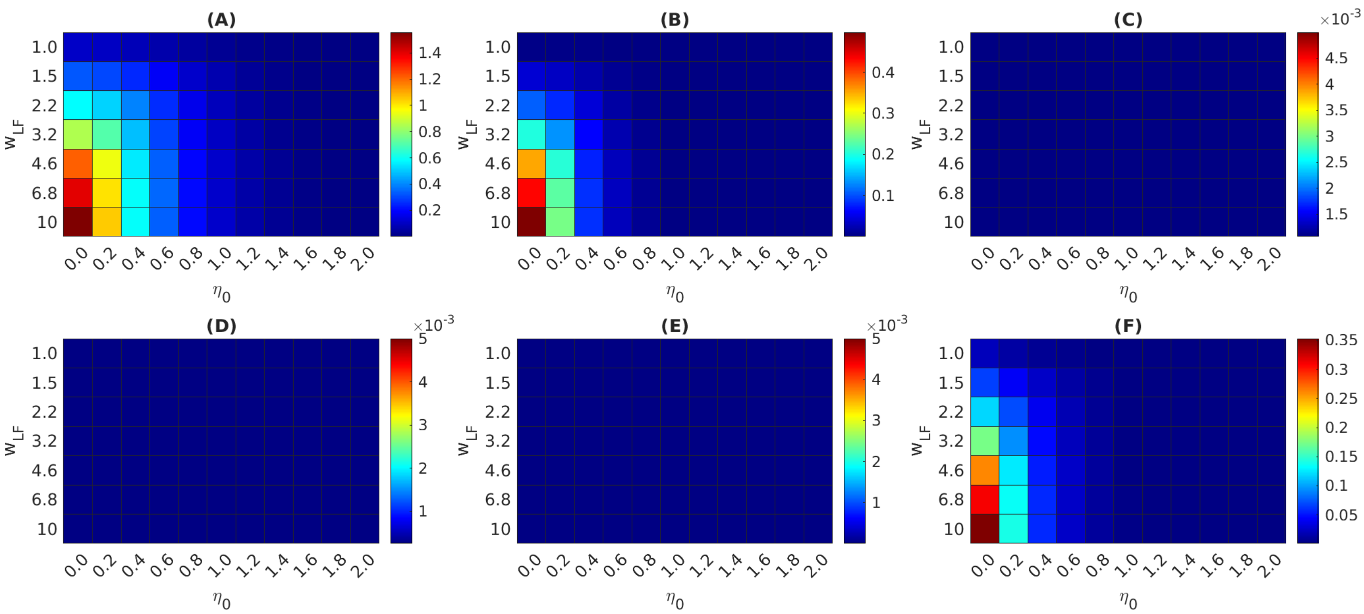

FIG. 13. (A) $\mathcal{I}_{\mathrm{L}\to\mathrm{F}}$, (B)$\sigma_{\mathrm{L}\to\mathrm{F}}$, (C)$\mathcal{S}_{\mathrm{L}\to\mathrm{F}}$, (D) $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$, (E)$\sigma_{\mathrm{F}\to\mathrm{L}}$, (F)$\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$ for interaction type B in Fig. 1, only with the leader forgetting its past once every two time steps.(we may move this to SI)

$$\langle \boldsymbol{\theta}(t)\rangle_{R,\boldsymbol{w},\vec{\boldsymbol{r}}^t} = \tan^{-1}\left[\frac{\sum_{j:|\vec{r}_i^t-\vec{r}_j^t|\leq R}\left[w_{ii}\sin\theta_i^t + w_{ij}\sin\theta_j^t\right]}{\sum_{j:|\vec{r}_i^t-\vec{r}_j^t|\leq R}\left[w_{ii}\cos\theta_i^t + w_{ij}\cos\theta_j^t\right]}\right],$$

the derivation of which can be found in the appendix of [1].

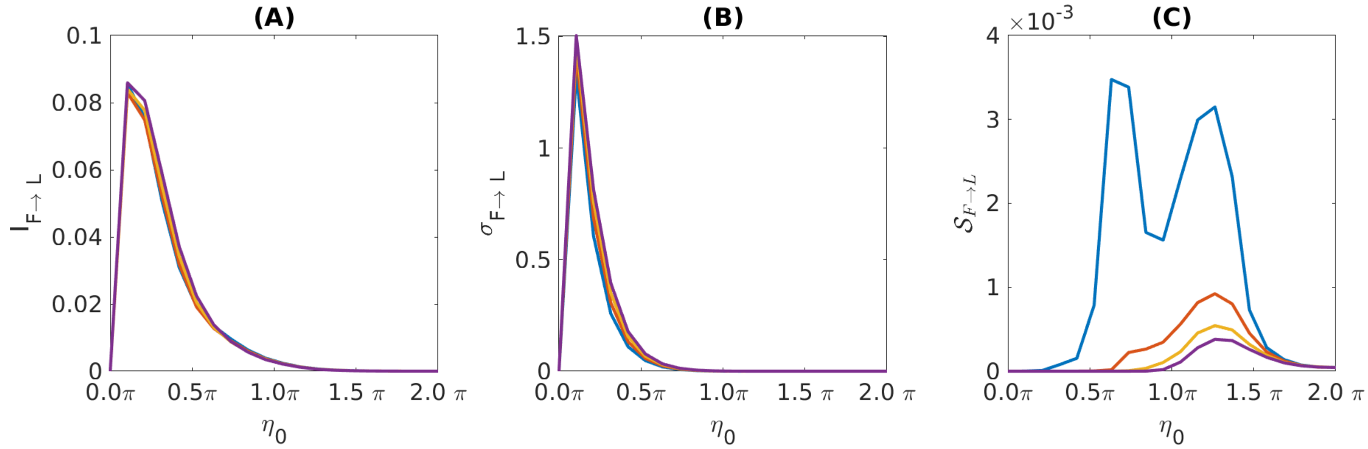

FIG. 14. Modes of information flow from follower to leader for the case where followers can influence followers for model A as a function of noise level η_0 with 2 agents (blue), 4 agents (red), 6 agents (orange) and 8 agents (purple). A) $\mathcal{I}_{\mathrm{F}\to\mathrm{L}}$, B) $\sigma_{\mathrm{F}\to\mathrm{L}}$, and C) $\mathcal{S}_{\mathrm{F}\to\mathrm{L}}$.

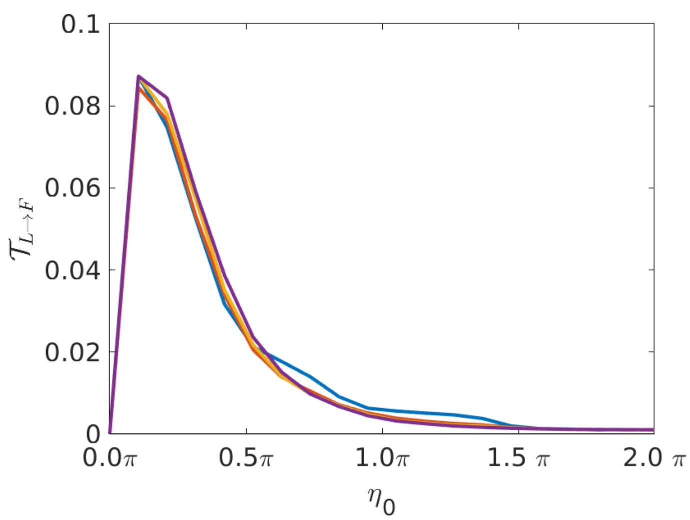

FIG. 15. $\mathcal{T}_{\mathrm{F}\to\mathrm{L}}$ as a function of noise level η_0 with 2 agents (blue), 4 agents (red), 6 agents (orange) and 8 agents (purple) for model A.

[1] Udoy S Basak, Sulimon Sattari, Kazuki Horikawa, and Tamiki Komatsuzaki. Inferring domain of interactions among particles from ensemble of trajectories. Physical Review E, 102(1):012404, 2020.

[2] Margret Bauer, John W Cox, Michelle H Caveness, James J Downs, and Nina F Thornhill. Finding the direction of disturbance propagation in a chemical process using transfer entropy. IEEE Transactions on Control Systems Technology, 15(1):12–21, 2006.

[3] Steven L Bressler and Anil K Seth. Wiener–granger causality: a well established methodology. Neuroimage, 58(2):323–329, 2011.

[4] Sachit Butail, Violet Mwaffo, and Maurizio Porfiri. Model-free information-theoretic approach to infer leadership in pairs of zebrafish. Physical Review E, 93(4):042411, 2016.

[5] Sachit Butail, Violet Mwaffo, and Maurizio Porfiri. Model-free information-theoretic approach to infer leadership in pairs of zebrafish. Physical Review E, 93(4):042411, 2016.

[6] SA Campuzano, Angelo De Santis, Fco Javier Pavón-Carrasco, María Luisa Osete, and E Qamili. New perspectives in the study of the earth's magnetic field and climate connection: The use of transfer entropy. PloS One, 13(11):e0207270, 2018.

[7] T. M. Cover and J. A. Thomas. Elements of Information Theory. Wiley-Interscience, New York, second edition, 2006.

[8] Alfonso Delgado-Bonal, Alexander Marshak, Yuekui Yang, and Daniel Holdaway. Analyzing changes in the complexity of climate in the last four decades using merra-2 radiation data. Scientific Reports, 10(1):1–8, 2020.

[9] Joshua Garland, Andrew M Berdahl, Jie Sun, and Erik M Bollt. Anatomy of leadership in collective behaviour. Chaos: An Interdisciplinary Journal of Nonlinear Science, 28(7):075308, 2018.
[10] Tobias Gerken, Benjamin L Ruddell, Rong Yu, Paul C Stoy, and Darren T Drewry. Robust observations of land-to-atmosphere feedbacks using the information flows of fluxnet. NPJ Climate and Atmospheric Science, 2(1):1–10, 2019.
[11] Ryan G James, Blanca Daniella Mansante Ayala, Bahti Zakirov, and James P Crutchfield. Modes of information flow. arXiv preprint arXiv:1808.06723, 2018.
[12] Ryan G James, Nix Barnett, and James P Crutchfield. Information flows? a critique of transfer entropies. Physical Review Letters, 116(23):238701, 2016.
[13] Jaeseung Jeong, John C Gore, and Bradley S Peterson. Mutual information analysis of the eeg in patients with alzheimer's disease. Clinical Neurophysiology, 112(5):827–835, 2001.
[14] Warren M Lord, Jie Sun, Nicholas T Ouellette, and Erik M Bollt. Inference of causal information flow in collective animal behavior. IEEE Transactions on Molecular, Biological and Multi-Scale Communications, 2(1):107–116, 2016.
[15] Ueli M Maurer. Secret key agreement by public discussion from common information. IEEE Transactions on Information Theory, 39(3):733–742, 1993.
[16] Ueli M Maurer and Stefan Wolf. Unconditionally secure key agreement and the intrinsic conditional information. IEEE Transactions on Information Theory, 45(2):499–514, 1999.
[17] Violet Mwaffo, Sachit Butail, and Maurizio Porfiri. Analysis of pairwise interactions in a maximum likelihood sense to identify leaders in a group. Frontiers in Robotics and AI, 4:35, 2017.
[18] Violet Mwaffo, Jishnu Keshavan, Tyson Hedrick, and Sean Humbert. A data-driven method to dissect the dynamics of the causal influence in complex dynamical systems. In 2018 IEEE Workshop on Complexity in Engineering (COMPENG), pages 1–5. IEEE, 2018.
[19] Violet Mwaffo, Jishnu Keshavan, Tyson L Hedrick, and Sean Humbert. Detecting intermittent switching leadership in coupled dynamical systems. Scientific Reports, 8(1):1–20, 2018.
[20] Máté Nagy, Zsuzsa Ákos, Dora Biro, and Tamás Vicsek. Hierarchical group dynamics in pigeon flocks. Nature, 464(7290):890–893, 2010.
[21] Shinnosuke Nakayama, Manuel Ruiz Marín, Maximo Camacho, and Maurizio Porfiri. Plasticity in leader–follower roles in human teams. Scientific Reports, 7(1):1–9, 2017.
[22] N Orange and N Abaid. A transfer entropy analysis of leader-follower interactions in flying bats. The European Physical Journal Special Topics, 224(17):3279–3293, 2015.
[23] Maurizio Porfiri. Inferring causal relationships in zebrafish-robot interactions through transfer entropy: A small lure to catch a big fish. Animal Behavior and Cognition, 5(4):341–367, 2018.
[24] Maurizio Porfiri, Raghu Ram Sattanapalle, Shinnosuke Nakayama, James Macinko, and Rifat Sipahi. Media coverage and firearm acquisition in the aftermath of a mass shooting. Nature Human Behaviour, 3(9):913–921, 2019.
[25] Myriam Reffay, Maria-Carla Parrini, Olivier Cochet-Escartin, Benoıt Ladoux, Axel Buguin, S Coscoy, François Amblard, Jacques Camonis, and Pascal Silberzan. Interplay of rhoa and mechanical forces in collective cell migration driven by leader cells. Nature Cell Biology, 16(3):217–223, 2014.
[26] Myriam Reffay, Laurence Petitjean, Sylvie Coscoy, Erwan Grasland-Mongrain, Francois Amblard, Axel Buguin, and Pascal Silberzan. Orientation and polarity in collectively migrating cell structures: statics and dynamics. Biophysical Journal, 100(11):2566–2575, 2011.
[27] Thomas Schreiber. Measuring information transfer. Physical Review Letters, 85(2):461, 2000.
[28] Richard E Spinney, Mikhail Prokopenko, and Joseph T Lizier. Transfer entropy in continuous time, with applications to jump and neural spiking processes. Physical Review E, 95(3):032319, 2017.
[29] Ariana Strandburg-Peshkin, Damien R Farine, Iain D Couzin, and Margaret C Crofoot. Shared decision-making drives collective movement in wild baboons. Science, 348(6241):1358–1361, 2015.
[30] Jie Sun, Dane Taylor, and Erik M Bollt. Causal network inference by optimal causation entropy. SIAM Journal on Applied Dynamical Systems, 14(1):73–106, 2015.
[31] Colin J Torney, Myles Lamont, Leon Debell, Ryan J Angohiatok, Lisa-Marie Leclerc, and Andrew M Berdahl. Inferring the rules of social interaction in migrating caribou. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1746):20170385, 2018.
[32] Raul Vicente, Michael Wibral, Michael Lindner, and Gordon Pipa. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. Journal of Computational Neuroscience, 30(1):45–67, 2011.
[33] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. Physical Review Letters, 75(6):1226, 1995.
[34] Michael Wibral, Benjamin Rahm, Maria Rieder, Michael Lindner, Raul Vicente, and Jochen Kaiser. Transfer entropy in magnetoencephalographic data: quantifying information flow in cortical and cerebellar networks. Progress in Biophysics and Molecular Biology, 105(1-2):80–97, 2011.
[35] Michael Wibral, Raul Vicente, and Michael Lindner. Transfer entropy in neuroscience. In Directed Information Measures in Neuroscience, pages 3–36. Springer, 2014.
[36] Patricia Wollstadt, Mario Martínez-Zarzuela, Raul Vicente, Francisco J Díaz-Pernas, and Michael Wibral. Efficient transfer entropy analysis of non-stationary neural time series. PloS One, 9(7):e102833, 2014.
[37] Naoya Yamaguchi, Takeomi Mizutani, Kazushige Kawabata, and Hisashi Haga. Leader cells regulate collective cell migration via rac activation in the downstream signaling of integrin β1 and pi3k. Scientific Reports, 5(1):1–8, 2015.