2012 3rd International Workshop on Cognitive Incromation Processing (CIP)

Effective Information Flow over Mobile Adaptive Networks

Sheng-Yuan Tu and Ali H. Sayed

Department of Electrical Engineering
University of California, Los Angeles

***Abstract*—Collective motion is a remarkable phenomenon in biological systems. There have been several models in the literature to regenerate this type of motion, such as averaging consensus strategies where nodes continuously average the velocity vectors of their neighbors. While many models are able to generate forms of collective motion, they nevertheless neglect the important fact that the most informed nodes in a network tend to modulate their information into their speeds. In this work, we show how the speed information can be exploited and incorporated into the design of the combination rules for mobile networks. The analysis leads to a sigmoidal function construction, and the results show that the proposed combination rule leads to more effective information flow over networks of mobile agents.**

***Index Terms*—Self-organization, adaptive networks, diffusion adaptation, fish schools, collective motion, information flow.**

I. Introduction

Self-organization is observed in several physical and biological systems [1], [2]. For example, fish join together in schools to avoid attacks by predators and improve foraging efficiency; birds fly in V-formation; and bees swarm towards a new hive. In these cases, a global pattern of behavior emerges from localized interactions among the individual components of the network. In earlier works [3]–[6], we used the concept of diffusion adaptation to model and regenerate these kinds of complex behavior.

In earlier works [7], [8], it has been suggested that the patterns of collective motion observed in nature can be modeled by having each node move along the average direction of motion of its neighbors. However, recent experiments on the behavioral rules of fish schools appear to challenge this traditional averaging strategy [9].

We argue in this work that in order to improve information transfer over a network of interacting agents, nodes should give higher weights to their most informed neighbors. In general, nodes do not know beforehand which other nodes in their neighborhoods are more informed. However, in biological networks, the speed of motion of a node usually conveys information about how well informed it is. For example, fish move faster when they feel danger or sense food. In this paper, motivated by this observation, we incorporate speed into the design of the combination weights over networks. By doing so, we show that the weights will need to be chosen according to a sigmoidal rule. We also show that this design leads to an effective flow of information across the network at faster rates than other more conventional designs.

This work was supported in part by NSF grants CCF-1011918 and CCF-0942936. E-mail: {shinetu, sayed}@ee.ucla.edu

II. Diffusion Adaptation Strategy

A. Algorithm Description

Consider a collection of N nodes distributed over some region in space. Two nodes are said to be neighbors if they can share information. The set of neighbors of node k is denoted by $\mathcal{N}_k$. The nodes would like to estimate some unknown column vector, w°, of size M. At every time instant, i, each node k observes realizations $\{d_k(i), u_{k,i}\}$ of a scalar random process $\boldsymbol{d}_k(i)$ and a $1 \times M$ random process $\boldsymbol{u}_{k,i}$ with covariance matrix $R_{u,k} = \mathbb{E}\boldsymbol{u}_{k,i}^*\boldsymbol{u}_{k,i} > 0$. All vectors in our treatment are column vectors with the exception of the regression vector, $\boldsymbol{u}_{k,i}$. The random processes $\{\boldsymbol{d}_k(i), \boldsymbol{u}_{k,i}\}$ are assumed to be related to w° via a linear regression model of the form [10]:

$$\boldsymbol{d}_k(i) = \boldsymbol{u}_{k,i}w^\circ + \boldsymbol{n}_k(i) \tag{1}$$

where $\boldsymbol{n}_k(i)$ is measurement noise with variance $\sigma_{n,k}^2$ and assumed to be temporally white and spatially independent, i.e.,

$$\mathbb{E}\boldsymbol{n}_k^*(i)\boldsymbol{n}_l(j) = \sigma_{n,k}^2 \cdot \delta_{kl} \cdot \delta_{ij} \tag{2}$$

in terms of the Kronecker delta function. The noise $\boldsymbol{n}_k(i)$ is also assumed to be independent of $\boldsymbol{u}_{l,j}$ for all l and j. All random processes are assumed to be zero mean.

The objective of the network is to estimate w° in a distributed manner through an online learning process. The nodes estimate w° by seeking to minimize the following global cost function:

$$J^{\text{glob}}(w) \triangleq \sum_{k=1}^{N} \mathbb{E}|\boldsymbol{d}_k(i) - \boldsymbol{u}_{k,i}w|^2 \tag{3}$$

Several diffusion adaptation schemes for solving (3) in a distributed manner were proposed in [11], [12]. One such scheme is the Combine-then-Adapt (CTA) diffusion algorithm [12]. It operates as follows. We select an $N \times N$ left-stochastic matrix A with nonnegative entries $\{a_{l,k} \geq 0\}$ satisfying:

$$A^T\mathbb{1} = \mathbb{1} \text{ and } a_{l,k} = 0 \text{ if } l \notin \mathcal{N}_k \tag{4}$$

where $\mathbb{1}$ is a vector of size N with all entries equal to one. The entry $a_{l,k}$ denotes the weight on the link connecting node l to node k. The CTA algorithm consists of two steps. The first step

(5a) is a consultation (combination) step where the weight estimates $\{w_{l,i-1}\}$ from the neighborhood are combined through the weights $\{a_{l,k}\}$ to obtain the intermediate estimate $\psi_{k,i-1}$. The second step (5b) involves local adaptation, where node k uses its own data $\{d_k(i), u_{k,i}\}$ to update the estimate at node k from $\psi_{k,i-1}$ to $w_{k,i}$. The algorithm is described as follows:

$$\begin{cases} \psi_{k,i-1} = \sum_{l \in \mathcal{N}_k} a_{l,k} w_{l,i-1} & \text{(5a)} \\ w_{k,i} = \psi_{k,i-1} + \mu_k u_{k,i}^* [d_k(i) - u_{k,i}\psi_{k,i-1}] & \text{(5b)} \end{cases}$$

where μ_k is the positive step-size used by node k.

In biological networks, the behavior of the network is often influenced more heavily by a small fraction of informed agents as happens, for example, with bees and fish [13]–[15]. This observation motivates us to consider two types of nodes: informed nodes and uninformed nodes. The former receive new data $\{d_k(i), u_{k,i}\}$ regularly and perform both the consultation step (5a) and the adaptation step (5b), while uninformed nodes do not collect data and only participate in the consultation step (5a). To model uninformed nodes in the network, we set $\mu_k = 0$ if node k is uninformed.

B. Information Flow over Adaptive Networks

The performance of diffusion algorithms in the presence of uninformed nodes has been studied in [16], including questions related to how the mean-square-deviation (MSD) of the network varies as a function of the proportion of informed nodes. In the current work, we are interested in examining instead the rate of the information flow, i.e., how fast the network converges towards steady-state and how combination weights can be chosen to speed up the flow of information across the network. The convergence rate is denoted by r so that the smaller the value of r is, the faster the rate of convergence. It was shown in [10], [12], [16] that

$$\boxed{r = \left[\rho((I - \mathcal{MR})\mathcal{A}^T)\right]^2} \tag{6}$$

where $\rho(\cdot)$ denotes the spectral radius of its matrix argument and

$$\mathcal{A} = A \otimes I_M, \quad \mathcal{M} = \text{diag}\{\mu_k I_M\}, \quad \mathcal{R} = \text{diag}\{R_{u,k}\} \tag{7}$$

and where the operation $\otimes$ denotes the Kronecker product of two matrices. Note from (6) that r depends on the combination matrix A and on the spatial distribution of the informed nodes through the matrix $\mathcal{M}$. Under the assumption that $\mu_k = \mu$ for all informed nodes, $R_{u,k} = R_u$ for all k, and that the step-size is small enough such that

$$\mu\rho(R_u) < 1 \tag{8}$$

it can be shown that the convergence rate is bounded by [16]:

$$(1 - \mu\lambda_M(R_u))^2 \leq r < 1 \tag{9}$$

where $\lambda_M(R_u)$ is the smallest eigenvalue of R_u. We will show that by appropriately selecting the combination matrix A in any connected network (where a path always exists between any two arbitrary nodes), the convergence rate (6) can achieve the lower bound provided by (9), that is, the network is able to converge to steady-state at the fastest rate.

Let $\mathcal{N}_I$ denote the set of informed nodes and let N_I denote the number of informed nodes in the network. Without loss of generality, we assume the first N_I nodes are informed, i.e., $\mathcal{N}_I = \{1, \ldots, N_I\}$. The combination matrix A can be partitioned in the following manner:

$$A = \left[\begin{array}{c|c} A_{II} & A_{IU} \\ \hline A_{UI} & A_{UU} \end{array}\right] \tag{10}$$

where the sub-matrices A_{II} and A_{UU} have size $N_I \times N_I$ and $(N - N_I) \times (N - N_I)$, respectively. Thus, the matrix A_{II} collects the weights among the informed nodes and A_{UI} collects the weights from uninformed to informed nodes; likewise for $\{A_{UU}, A_{IU}\}$. The matrix $(I - \mathcal{MR})\mathcal{A}^T$ can then be written as:

$$(I - \mathcal{MR})\mathcal{A}^T = \begin{bmatrix} A_{II}^T \otimes (I_M - \mu R_u) & A_{UI}^T \otimes (I_M - \mu R_u) \\ A_{IU}^T \otimes I_M & A_{UU}^T \otimes I_M \end{bmatrix} \tag{11}$$

The following result gives a condition on A so that the convergence rate achieves its lower bound.

Lemma 1. *For any connected CTA network (5) with at least one informed node, if the sub-matrices $\{A_{UI}, A_{UU}\}$ of the combination matrix A in (10) satisfy:*

$$A_{UI} = 0, \quad \rho(A_{UU}) \leq 1 - \mu\lambda_M(R_u) \tag{12}$$

then the convergence rate r in (6) achieves its lower bound, i.e., $r = (1 - \mu\lambda_M(R_u))^2$.

Proof: Since $A_{UI} = 0$, the matrix $(I - \mathcal{MR})\mathcal{A}^T$ in (11) becomes lower block triangular, and its spectral radius is the maximum of $\rho(A_{II}^T \otimes (I_M - \mu R_u))$ and $\rho(A_{UU}^T \otimes I_M)$. Moreover, since $A_{UI} = 0$ and using (4), we conclude that A_{II} becomes left-stochastic. Hence, $\rho(A_{II}) = 1$ and it follows that

$$\begin{aligned} \rho(A_{II}^T \otimes (I_M - \mu R_u)) &= \rho(A_{II}^T) \cdot \rho(I_M - \mu R_u) \\ &= 1 - \mu\lambda_M(R_u) \end{aligned} \tag{13}$$

where we used assumption (8). Moreover, since $\rho(A_{UU}) \leq 1 - \mu\lambda_M(R_u)$, we get

$$\rho(A_{UU}^T \otimes I_M) = \rho(A_{UU}^T) \cdot \rho(I_M) \leq 1 - \mu\lambda_M(R_u) \tag{14}$$

Then, $r = (1 - \mu\lambda_M(R_u))^2$. ∎

Theorem 1. *For any connected CTA network (5) with at least one informed node, there exists a combination matrix A such that the network achieves the fastest convergence rate.*

Proof: From Lemma 1, it suffices to show that we are always able to construct a combination matrix A satisfying (12). First, we index the nodes such that the smaller the distance (number of hops) from a node to the set $\mathcal{N}_I$ is, the smaller the index of the node is. This can be done by first indexing informed nodes in any order, and then indexing the uninformed nodes next to the informed nodes in any order, and so on (see the middle plot of Fig. 1). Second, besides

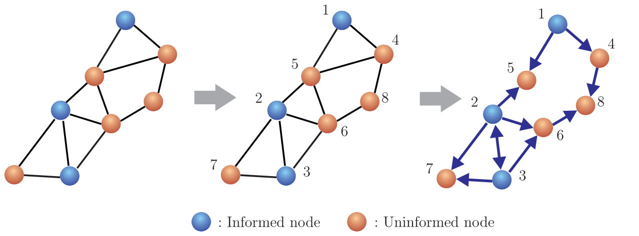

Fig. 1. An illustration of a connected network with three informed nodes (left) and one way to achieve the fastest convergence rate (right).

condition (4), we further require the weights $\{a_{l,k}\}$ to satisfy the following rule:

$$\begin{cases} \sum\limits_{l\in\mathcal{N}_I\cap\mathcal{N}_k} a_{l,k} = 1, & \text{if } \mathcal{N}_I\cap\mathcal{N}_k \neq \emptyset \quad (15a) \\ \sum\limits_{l<k} a_{l,k} = 1, & \text{otherwise} \quad (15b) \end{cases}$$

That is, if there are informed nodes in the neighborhood of node k, then it will assign positive combination weights to those nodes only; otherwise, node k will assign positive combination weights to neighbors with lower indices than k (i.e., those closer to informed nodes). The example in Fig. 1 leads to a matrix A^T of the form below, where the directions of the arrows in the right plot of Fig. 1 indicate the allowed direction of information flow, i.e., the combination weights in the reverse directions are zero:

$$A^T = \left[\begin{array}{ccc|ccccc} 1 & 0 & 0 & & & & & \\ 0 & a & 1-a & & & & & \\ 0 & b & 1-b & & & & & \\ \hline 1 & 0 & 0 & 0 & & & & \\ c & 1-c & 0 & 0 & 0 & & & \\ 0 & d & 1-d & 0 & 0 & 0 & & \\ 0 & e & 1-e & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & f & 0 & 1-f & 0 & 0 \end{array}\right] \quad (16)$$

with $a, b, c, d, e, f \in [0, 1]$. Since the weight arriving at an informed node from an uninformed node is always zero, $A_{UI} = 0$. In addition, for an uninformed node k, the weight $a_{l,k}$ is equal to zero if $l \geq k$. Then, the matrix A_{UU}^T in (10) is a lower triangular matrix with zero diagonal entries and $\rho(A_{UU}) = 0 < 1 - \mu\lambda_M(R_u)$. ■

III. Mechanism for Collective Motion

In this section, we first describe a mechanism for mobile nodes to perform collective motion in the plane, and later examine the flow of information through the resulting network. At time i, node k is at location vector, $x_{k,i}$, and moves at velocity, $v_{k,i}$. Node k is able to observe the locations and velocities of its neighbors, i.e., $\{x_{l,i}, v_{l,i}\}$ for $l \in \mathcal{N}_k$. Node k updates its location according to the rule:

$$x_{k,i+1} = x_{k,i} + \triangle t \cdot v_{k,i+1} \quad (17)$$

where $\triangle t$ is the time step. Several factors influence the update of the velocity vector from $v_{k,i}$ to $v_{k,i+1}$ such as the desire to move towards a food source or away from danger, the desire to move in coordination with the other nodes, and the desire to avoid collisions by keeping a safe distance from neighbors. The determination of $v_{k,i+1}$ for node k should depend only on its observations, i.e., $\{x_{l,i}, v_{l,i}\}$ for $l \in \mathcal{N}_k$. Based on the results from [3], we assume that nodes adjust their velocity according to the following CTA diffusion algorithm:

$$\boxed{\begin{aligned} \psi_{k,i} &= \sum_{l\in\mathcal{N}_k} a_{l,k} v_{l,i} \\ v_{k,i+1} &= \psi_{k,i} + \mu_k(v^d_{k,i+1} - \psi_{k,i}) + \gamma_k \delta_{k,i} \end{aligned}} \quad (18)$$

where γ_k is a non-negative scalar. Expression (18) involves three components. The first component in (18) is determined by the weighted velocity $\psi_{k,i}$ in the neighborhood of node k (i.e., the combination step). This component helps the nodes move coherently [8]. The second component in (18) corresponds to a desired velocity vector, $v^d_{k,i+1}$. This direction may be induced by the sensing by node k of the location of food or danger (e.g., a predator). The third component in (18) involves the term $\delta_{k,i}$, which is defined as [3]:

$$\boxed{\delta_{k,i} = \sum_{l\in\mathcal{N}_k\backslash\{k\}} b_{l,k} \left(\|x_{l,i} - x_{k,i}\| - r\right) \frac{x_{l,i} - x_{k,i}}{\|x_{l,i} - x_{k,i}\|}} \quad (19)$$

with a positive parameter r and non-negative weights $\{b_{l,k}\}$ satisfying $b_{k,k} = 0$ and

$$\sum_{l\in\mathcal{N}_k\backslash\{k\}} b_{l,k} = 1 \text{ and } b_{l,k} = 1 \text{ if } l \notin \mathcal{N}_k \quad (20)$$

The term $\delta_{k,i}$ allows node k to maintain a distance r from its neighbors. Compared to (5), we find that $\delta_{k,i}$ in (18) is an extra term. Nevertheless, as shown in [3], the term $\delta_{k,i}$ will be close to zero as time evolves. In the next section, we derive a combination rule to select the weights $\{a_{l,k}, b_{l,k}\}$.

IV. Combination Rules

The choice of the combination weights $\{a_{l,k}, b_{l,k}\}$ in (18) and (19) influences the way the nodes interact with each other. Different choices for the weights not only lead to different patterns of behavior, but they also influence the flow of information through the network, as revealed by (6). In earlier works [7], [8], the uniform combination rule (or averaging strategy) has been employed to regenerate the collective motion exhibited by fish schooling or bird flocking. That is, the weights were set to

$$a_{l,k} = b_{l,k} = \begin{cases} 1/(n_k - 1), & \text{if } l \in \mathcal{N}_k \setminus \{k\} \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

if node k is uninformed, where n_k denotes the size of the neighborhood of node k (or its degree). If node k is informed, the weights $\{b_{l,k}\}$ remain the same while the weights $\{a_{l,k}\}$ change to the rule employed in [17], [18]:

$$a_{l,k} = \begin{cases} \alpha/(\alpha + n_k - 1), & \text{if } l = k \\ 1/(\alpha + n_k - 1), & \text{if } l \in \mathcal{N}_k \setminus \{k\} \\ 0, & \text{otherwise} \end{cases} \quad (22)$$

where α is a positive weighting factor.

In this paper, we would like to select the combination weights such that the information is transferred through the network in the most efficient way (i.e., at the fastest convergence rate). As suggested by Theorem 1, this can be accomplished if every node employs rule (15). However, in general, nodes do not know whether their neighbors are informed or not, and the spatial distribution of informed nodes. Nevertheless, the speed of a node usually reflects the quality of its private information. For example, fish tend to move faster when they sense food or feel danger. Thus, nodes may use the speed of their neighbors to infer information about how informed they are. To explain this idea, we drop the time index, denote the velocity for node k in the $2D-$plane by $v_k = (v_{1,k}, v_{2,k})$, and let s_k denote its speed:

$$s_k = \|v_k\| = \sqrt{v_{1,k}^2 + v_{2,k}^2} \tag{23}$$

In addition, let $\mathbb{I}_k$ be an indicator function for node k whose value is equal to 1 if node k is informed; otherwise the value is equal to 0. Then, the combination weights $\{a_{l,k}, b_{l,k}\}$ can be selected in proportion to the probability that node l is informed given its speed s_l:

$$a_{l,k} \propto \Pr(\mathbb{I}_l = 1|s_l) \tag{24}$$

and likewise for $\{b_{l,k}\}$. Note that since node k knows whether it possesses information or not, the probability $\Pr(\mathbb{I}_k = 1|s_k)$ is simply zero or one. In this way, a node essentially places positive weights only to neighboring informed nodes, as in (15). In the following, we give a model for the velocity vector v_k and evaluate the probability $\Pr(\mathbb{I}_l = 1|s_l)$.

To begin with, in the absence of neighbors, we assume that the speed of any node k is set as follows:

$$s^\circ = \begin{cases} c_0, & \text{if } \mathbb{I}_k = 0 \\ c_1, & \text{if } \mathbb{I}_k = 1 \end{cases} \tag{25}$$

with $c_1 > c_0$. In this way, when neighbors are not present, node k moves at speed c_0 when it is uninformed; otherwise it moves faster at speed c_1 towards a food source or away from danger (i.e., $\|v_k^d\| = c_1$ in (18)). However, when neighbors are present, the motion of node k will be affected by its neighbors according to (18); in this case, node k will not move at a constant speed (c_0 or c_1). To take this effect into account, we introduce the following model:

$$\boldsymbol{v}_{1,k} = s^\circ \cos\boldsymbol{\theta}_k + \boldsymbol{n}_{1,k} \tag{26}$$

$$\boldsymbol{v}_{2,k} = s^\circ \sin\boldsymbol{\theta}_k + \boldsymbol{n}_{2,k} \tag{27}$$

where $\boldsymbol{\theta}_k$ is the moving direction of node k, and $\boldsymbol{n}_{1,k}$ and $\boldsymbol{n}_{2,k}$ are Gaussian random variables with zero mean and variance σ_n^2. Moreover, $\boldsymbol{\theta}_k$, $\boldsymbol{n}_{1,k}$, and $\boldsymbol{n}_{2,k}$ are assumed to be independent of each other. Expressions (26)-(27) model the perturbation caused by the neighbors of node k. Therefore, the velocity vector v_k, given $\boldsymbol{\theta}_k = \theta_k$, becomes a Gaussian random vector with mean $\begin{bmatrix} s^\circ\cos\theta_k & s^\circ\sin\theta_k\end{bmatrix}^T$ and covariance matrix $\sigma_n^2 I_2$. Thus, the speed s_k in (23) is a Rician random variable with parameters $\{s^\circ, \sigma_n^2\}$ [19] and the probability density function (pdf) of s_k, given $\boldsymbol{\theta}_k = \theta_k$, can be written as:

$$f(s_k|s^\circ, \sigma_n^2, \boldsymbol{\theta}_k = \theta_k) = \frac{s_k}{\sigma_n^2}\exp\left[\frac{-(s_k^2 + s^{\circ 2})}{2\sigma_n^2}\right] I_0\left(\frac{s_k s^\circ}{\sigma_n^2}\right) \tag{28}$$

where $I_0(z)$ is the modified Bessel function of the first kind with order zero, or

$$I_0(z) = \sum_{m=0}^{\infty}\left[\frac{(z/2)^m}{m!}\right]^2 \tag{29}$$

Note that expression (28) is independent of θ_k. Then, the pdf $f(s_k|s^\circ, \sigma_n^2)$ is identical to (28).

Using Bayes' rule, the probability $\Pr(\mathbb{I}_l = 1|s_l)$ can be evaluated by

$$\Pr(\mathbb{I}_l = 1|s_l) = \frac{\Pr(\mathbb{I}_l = 1) f(s_l|\mathbb{I}_l = 1)}{\sum_{m=0}^{1}\Pr(\mathbb{I}_l = m) f(s_l|\mathbb{I}_l = m)} \tag{30}$$

where

$$f(s_l|\mathbb{I}_l = m) = f(s_l|s^\circ = c_m, \sigma_n^2) \tag{31}$$

Since nodes do not have prior information about whether other nodes are informed or not, they simply set the prior probabilities, $\Pr(\mathbb{I}_l = 1)$ and $\Pr(\mathbb{I}_l = 0)$, to equal values (namely, 1/2). Substituting the pdf from (28) into (30), we arrive at

$$\Pr(\mathbb{I}_l = 1|s_l) = \left[1 + \exp\left(\frac{c_1^2 - c_0^2}{2\sigma_n^2}\right)\frac{I_0\left(\frac{s_l c_0}{\sigma_n^2}\right)}{I_0\left(\frac{s_l c_1}{\sigma_n^2}\right)}\right]^{-1} \tag{32}$$

However, the Bessel function (29) is difficult to compute. We can use the following approximation

$$\frac{I_0\left(\frac{s_l c_0}{\sigma_n^2}\right)}{I_0\left(\frac{s_l c_1}{\sigma_n^2}\right)} \approx \exp\left[\frac{-s_l(c_1 - c_0)}{\sigma_n^2}\right] \tag{33}$$

and expression (32) simplifies to

$$\Pr(\mathbb{I}_l = 1|s_l) \approx \left\{1 + \exp\left[\frac{c_1 - c_0}{\sigma_n^2}\left(\frac{c_1 + c_0}{2} - s_l\right)\right]\right\}^{-1} \tag{34}$$

Expressions (32) and (34) are depicted in Fig. 2 with parameters $(c_1, c_0, \sigma_n^2) = (4, 1, 1)$. We observe that the two curves are close to each other. Note that expression (34) admits a physical interpretation. The probability (34) attains the value of 0.5 when $s_l = (c_1 + c_0)/2$, which is the middle point of c_0 and c_1. That is, when the speed of one node passes the middle point, it has higher probability of being informed. In addition, the slope of the curve near the middle point is determined by $(c_1 - c_0)/\sigma_n^2$. If the difference between the speed of informed and uninformed nodes is large, nodes have better ability of distinguishing whether other nodes are informed or not. We assume that every node knows the values of the parameters (c_1, c_0, σ_n^2).

We conclude from (24) and Fig. 2 that the resulting combination rule exhibits a sigmoidal shape so that a node places

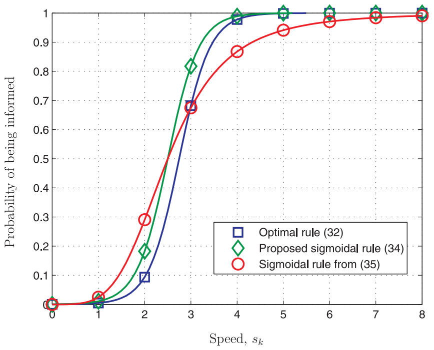

Fig. 2. Sigmoidal combination rules: the larger the speed of a node, the larger the weight assigned to it.

higher weights on faster-moving nodes. In this way, when a fast-moving node takes a sharp turn, for example, the effect of this behavior will ripple through the network at a faster rate and the remaining nodes will follow suit. The choice of the combination rule (34) is similar to the decision-making process in animal groups. When a node makes a decision (such as the decision to turn around and start moving in the opposite direction), the probability of other nodes following suit increases if the number of neighbors making a similar decision increases. Specifically, when the number of neighbors making the same decision passes some threshold value, the probability of other nodes following suit increases rapidly (i.e. sigmoidal type behavior). This phenomenon is called *quorum response* in animal group behavior [20]. Motivated by [20], another way to determine the probability $\Pr(\mathbb{I}_l = 1|s_l)$ is

$$\Pr(\mathbb{I}_l = 1|s_l) = \left[1 + \left(\frac{s_l}{K_1}\right)^{K_2}\right]^{-1} \quad (35)$$

Expression (35) is also shown in Fig. 2 for $(K1, K2) = (2.5, 4)$. Similar to (34), expression (35) attains the value of 0.5 when $s_l = K_1$ and the slope of the curve near the middle point is determined by K_2. We compare the performance of the two sigmoidal rules (34) and (35) in the next section.

V. Information Transfer in the Presence of Danger

In biological systems, the motion of the network is often influenced heavily by a small fraction of informed nodes. The experiment performed in Fig. 1 of [21] serves as a good example. In that experiment, when a few nodes on the boundary of the perimeter were frightened, these nodes rapidly changed their motion and reversed their orientation. The behavior propagated through the network very quickly. After a short period of time, the entire network ended up moving in the opposite direction relative to the original motion. We examine this effect and compare different combination rules.

We consider three combination rules: two sigmoidal rules from (34) with $(c_1, c_0, \sigma_n^2) = (4, 1, 1)$ and from (35) with $(K1, K2) = (2.5, 4)$, and the uniform rule (21) and (22) with $\alpha = 5$. The step-sizes are set to $\mu_k = 0.6$ for informed

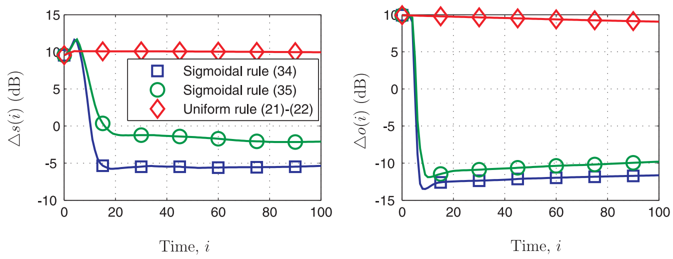

Fig. 4. Magnitude, $\triangle s(i)$, and orientation, $\triangle o(i)$, of the velocity of the center of mass relative the the trigger velocity.

nodes. Figure 3 shows simulation results for a network with $N = 100$ nodes. Initially, the velocities of the nodes are set to $v_k = (1, 0)$ for all k. To choose N_I threatened (informed) nodes, we first pick up the node with the largest x-coordinate and then choose $N_I - 1$ nodes that are closest to the chosen node. In simulations, we set $N_I = 2$. The desired velocities of the informed nodes are set to $v_{k,i}^d = v^d = (-4, 0)$ for 5 time steps with $\triangle t = 0.1$ sec. The resulting maneuver of the networks using the sigmoidal and uniform combination rules are shown in Fig. 3. The dots denote the locations of the nodes and their moving directions are indicated by the lines. Moreover, the nodes moving towards the positive (negative) x-direction are shown in red (blue). We observe that the motion of the informed nodes propagates rapidly through the entire network if the network employs the sigmoidal combination rules (34) or (35), while the network using the uniform combination rule (21) fails to transfer the motion through the entire network. To compare these three combination rules quantitatively, we measure the magnitude and orientation of the velocity of the center of mass of the network, v_i^g, relative to the desired velocity, v^d. That is, we introduce

$$\triangle s(i) = (\|v_i^g\| - \|v^d\|)^2 \quad (36)$$

$$\triangle o(i) = [\angle(v_i^g) - \angle(v^d)]^2 \quad (37)$$

where $v_i^g \triangleq \frac{1}{N}\sum_{k=1}^{N} v_{k,i}$. These two quantities are averaged over 100 experiments and shown in Fig. 4. We observe that the desired velocity of the informed nodes is successfully transferred through the network if the network adopts the sigmoidal combination rules. Moreover, comparing the two sigmoidal rules, we observe that rule (34) outperforms rule (35). *The results indicate that if the information of the nodes is modulated according to their speed, this mechanism improves the efficiency of information transfer over the network.*

VI. Conclusion

In this paper, we studied the interaction mechanism among agents in a mobile adaptive network. Even though the uniform combination rule is able to regenerate collective motion in biological systems, the sigmoidal combination rules show advantages in transferring information over the network.

References

[1] S. Camazine, J. L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz, and E. Bonabeau, *Self-Organization in Biological Systems*. Princeton University Press, 2003.

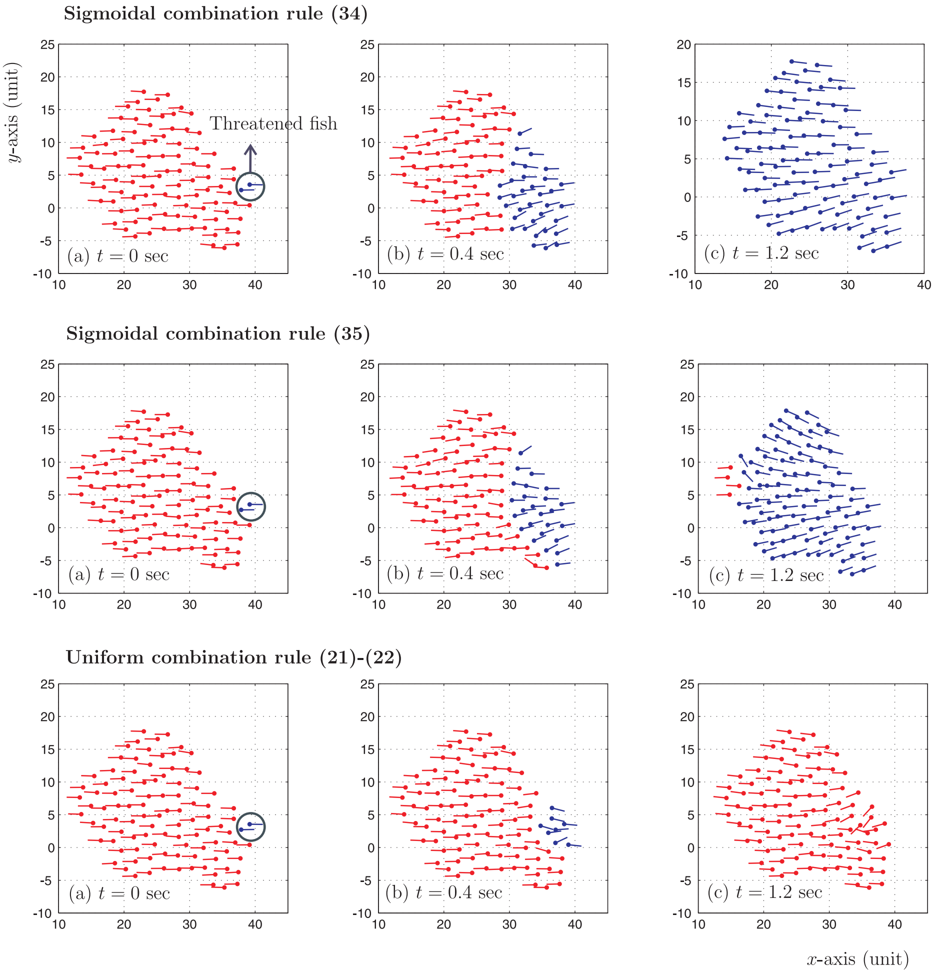

Fig. 3. Maneuver of fish schools with the two sigmoidal (top (34) and middle (35)) and uniform (bottom (21)-(22)) combination rules over time (a) $t = 0$ (b) $t = 0.4$ (c) $t = 1.2$ sec.

[2] I. D. Couzin, "Collective cognition in animal groups," *Trends in Cognitive Sciences*, vol. 13, pp. 36–43, Jan. 2009.

[3] S. Y. Tu and A. H. Sayed, "Mobile adaptive networks," *IEEE J. Selected Topics on Signal Processing*, vol. 5, no. 4, pp. 649–664, Aug. 2011.

[4] ——, "Cooperative prey herding based on diffusion adaptation," *Proc. IEEE ICASSP*, pp. 3752–3755, Prague, Czech Republic, May 2011.

[5] F. Cattivelli and A. H. Sayed, "Modeling bird flight formations using diffusion adaptation," *IEEE Trans. on Signal Processing*, vol. 59, no. 5, pp. 2038–2051, May 2011.

[6] J. Li and A. H. Sayed, "Modeling bee swarming behavior through diffusion adaptation with asymmetric information sharing," *EURASIP J. Advances Signal Process.*, pp. 1–14, doi:10.1186/1687-6180-2012-18, Jan. 2012.

[7] I. D. Couzin, J. Krause, R. James, G. D. Ruxton, and N. R. Franks, "Collective memory and spatial sorting in animal groups," *Journal of Theoretical Biology*, vol. 218, pp. 1–11, 2002.

[8] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups for mobile autonomous agents using nearest neighbor rules," *IEEE Trans. on Automatic Control*, vol. 48, pp. 988–1001, Jun. 2003.

[9] Y. Katz, C. C. Ioannou, K. Tunstrom, C. Huepe, and I. D. Couzin, "Inferring the structure and dynamics of interactions in schooling fish," *PNAS*, 2011.

[10] A. H. Sayed, *Adaptive Filters*. NJ. Wiley, 2008.

[11] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. on Signal Processing*, vol. 56, no. 7, pp. 3122–3136, Jul. 2008.

[12] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. on Signal Processing*, vol. 58, no. 3, pp. 1035–1048, Mar. 2010.

[13] A. Avitabile, R. A. Morse, and R. Boch, "Swarming honey bees guided by pheromones," *Ann. Entomol. Soc. Am.*, vol. 68, pp. 1079–1082, 1975.

[14] T. D. Seeley, R. A. Morse, and P. K. Visscher, "The natural history of the flight of honey bee swarms," *Psyche*, vol. 86, pp. 103–114, 1979.

[15] S. G. Reebs, "Can a minority of informed leaders determine the foraging movements of a fish shoal?" *Anim. Behav.*, vol. 59, pp. 403–409, 2000.

[16] S. Y. Tu and A. H. Sayed, "On the effects of topology and node distribution on learning over complex adaptive networks," *Proc. Asilomar Conference on Signals, Systems, and Computers*, pp. 1–5, Pacific Grove, CA, Nov. 2011.

[17] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, "Effective leadership and decision making in animal groups on the move," *Nature*, vol. 433, pp. 513–516, 2005.

[18] I. D. Couzin, C. C. Ioannou, G. Demirel, T. Gross, C. J. Torney, A. Hartnett, L. Conradt, S. A. Levin, and N. E. Leonard, "Uninformed individuals promote democratic consensus in animal groups," *Science*, vol. 332, pp. 1578–1580, 2011.

[19] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 2002.

[20] D. J. T. Sumpter and S. C. Pratt, "Quorum responses and consensus decision making," *Phil. Trans. R. Soc. B*, vol. 364, pp. 743–753, Dec. 2009.

[21] D. J. T. Sumpter, J. Buhl, D. Biro, and I. D. Couzin, "Information transfer in moving animal groups," *Theory in Biosciences*, vol. 127, no. 2, pp. 177–186, 2008.