

# Quantifying and Tracing Information Cascades in Swarms

X. Rosalind Wang<sup>1\*</sup>, Jennifer M. Miller<sup>2</sup>, Joseph T. Lizier<sup>1,3</sup>, Mikhail Prokopenko<sup>1</sup>, Louis F. Rossi<sup>2</sup>

**1** CSIRO Information and Communication Technologies Centre, Marsfield, New South Wales, Australia, **2** Department of Mathematical Sciences, University of Delaware, Newark, Delaware, United States of America, **3** Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

## Abstract

We propose a novel, information-theoretic, characterisation of cascades within the spatiotemporal dynamics of swarms, explicitly measuring the extent of collective communications. This is complemented by dynamic tracing of collective memory, as another element of distributed computation, which represents capacity for swarm coherence. The approach deals with both global and local information dynamics, ultimately discovering diverse ways in which an individual's spatial position is related to its information processing role. It also allows us to contrast cascades that propagate conflicting information with waves of coordinated motion. Most importantly, our simulation experiments provide the first direct information-theoretic evidence (verified in a simulation setting) for the long-held conjecture that the information cascades occur in waves rippling through the swarm. Our experiments also exemplify how features of swarm dynamics, such as cascades' wavefronts, can be filtered and predicted. We observed that maximal information transfer tends to follow the stage with maximal collective memory, and principles like this may be generalised in wider biological and social contexts.

**Citation:** Wang XR, Miller JM, Lizier JT, Prokopenko M, Rossi LF (2012) Quantifying and Tracing Information Cascades in Swarms. PLoS ONE 7(7): e40084. doi:10.1371/journal.pone.0040084

**Editor:** Gonzalo G. de Polavieja, Cajal Institute, Consejo Superior de Investigaciones Científicas, Spain

**Received:** January 21, 2012; **Accepted:** May 31, 2012; **Published:** July 12, 2012

**Copyright:** © 2012 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** LFR and JMM are funded by United States National Science Foundation grants CCF-0726556 and CCF-0829748. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No other external fundings were received for this study.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: Rosalind.Wang@csiro.au

## Introduction

Animal groups in nature often exhibit striking examples of spatial aggregation, e.g. schools of fish, swarms of locusts, herds of wildebeest, and flocks of birds [1–3]. Such aggregations may provide individuals with protection, mate choices, foraging, habitat assessment, migratory routes, etc. [4,5]. Complex large-scale patterns and structures emerge within swarms through individual decisions based on perception of local conditions. It has been observed that small perturbations cascade through an entire swarm in a wave-like manner [6], with these cascades conjectured to embody information transfer [7]. Even a few individuals may strongly bias the motion of an entire group. For instance, if a certain number of fish in close proximity turn together, this may result in a wave of turning across the whole group [8]. Formation of waves is a widespread phenomenon observed in animal groups [6,7,9], seeming to rapidly transfer information over long ranges. Such waves are typically conjectured as information cascades [7], and we aim to quantify these cascades in precise information-theoretic terms.

In a seminal work, Bikhchandani et al. [10] defined an “informational cascade” as a phenomenon occurring “when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behaviour of the preceding individual without regard to his own information”, i.e. via an independence of an individual's action from their private information signal. They identified two social regularities that can be explained by informational cascades: localised conformity of behaviour and fragility of mass behaviours. Their approach was not information-

theoretic and did not quantify a precise information content stored/acquired within a group or transferred by a cascade.

Information cascades in collective systems often result in a rapid autocatalytic adaptive response to changing conditions [7]. This heightened response allows the group to be extremely sensitive to weak or ambiguous external stimuli, though retaining some susceptibility to noise, incorrect decisions and false alarms [5,7,11].

Dall et al. [12] mentioned that public information favours group cohesion, argued that information implies utility as well as uncertainty reduction, and proposed an explicit statistical decision theory framework. Their approach did not quantify either the degree of swarm cohesiveness due to public information, or information cascades *per se*. They pointed out that Shannon-Weaver entropy and similar ideas focused on simple reductions of uncertainty do not suffice in organismal biology. We argue that the *information dynamics* model used in our study goes beyond these simple ideas by utilising a directed measure (transfer entropy [13]) for information cascades, as well as localising average information-theoretic quantities.

As pointed out by Katz *et al.* [14] important questions are how animals integrate information from widely disparate sources in real time [15] and how this nonlinear integration translates into higher-order collective computational capabilities. There is an emerging understanding that information is a crucial currency for animals from both a behavioural and evolutionary perspective [12,16]. In this work, we take an information-theoretic viewpoint on distributed computation occurring within swarms, utilising a recently introduced framework for local information dynamics.

Coherence in the swarm is ultimately related to collective *memory* (e.g. long range interaction) [15,17], which benefits individuals locally and the “localised conformity of behaviour” becomes efficient [10]. Thus our first hypothesis is that the collective memory within the swarm that is used for computation is captured by *Active Information Storage* (AIS) [18,19]. The local AIS of an agent in the system is the amount of information in its past that is used in predicting its next state. The overall swarm’s AIS is the average over all individuals at any given time.

Information cascades, on the other hand, are manifestations of long range *communications* that either dynamically reorganise the swarm reducing the “fragility of mass behaviour” [10] or propagate incorrect decisions [11]. Our second hypothesis is that information cascades are captured by conditional *Transfer Entropy* (TE) [20,21], which characterises the communication aspect of distributed computation. The local information transfer between a source and a destination agent is defined as the information provided by the source about the destination’s next state that was not contained in the past of the destination [13]. Importantly, TE properly measures a directed, dynamic transfer of information.

At this stage we would like to contrast the measures of *transfer entropy* and *information flow*. These measures must be used separately to quantify information transfer and causal information flow respectively.

Transfer entropy was introduced by Schreiber [13] and has been applied in different settings. For instance, in computational neuroscience, the study [22] presented a novel method for interregional connectivity analysis, using multivariate extensions to the mutual information and transfer entropy. The method identified the underlying directed information structure between brain regions, highlighting changes in the structure according to behavioral conditions. The study also pointed out differences between transfer entropy and Granger causality. The main advantage is the capture of nonlinear relationships because nonlinear coupling cannot be detected by linear methods (e.g. Granger causality, nor with the non-directional mutual information).

Other relevant neuroscientific studies include the work of Wibral *et al.* [23] which utilized transfer entropy analysis of magnetoencephalography (MEG) source-level signals in detecting changes in cortical and subcortical networks between the different auditory task types, the work of Chicharro and Ledberg [24] which considers brain as a biological system consisting of multiple interacting subsystems and shows that the influence of causal connections on the natural dynamics of the system often cannot be analysed in terms of the causal effect of one subsystem on another.

*Information flow* was proposed as a measure for causal information flow by Ay and Polani [25], and it is important to realise a crucial difference between (1) transfer entropy and (2) information flow. As argued by Lizier and Prokopenko [26], predictive transfer (measured with transfer entropy) refers to the amount of information that a source variable adds to the next state of a destination variable; i.e. “If I know the state of the source, how much does that help to predict the state of the destination?”. On the other hand, causal effect (measured with information flow) refers to the extent to which the source variable has a direct influence on the next state of a destination variable, i.e. “If I change the state of the source, to what extent does that alter the state of the destination?”.

The difference between transfer entropy as a method to capture information transfer, and information flow as a measure to capture causal effect/flow, is very important and may cast observations in a different light. In this work, we stay completely within the

interpretation of predictive information transfer, and do not make any claims on detecting causal information flows.

Memory typically refers to the storage of information by an agent or process to be used in its future. It can be understood in a wider (collective/distributed) context, where stigmergy is used as a means to share information between agents via environment [18]. Grassé [27] introduced the term stigmergy (“previous work directs and triggers new building actions”) to describe a decentralised pathway of information flow in social insects. Stigmergy is a mechanism of indirect coordination among agents acting in the environment, where local traces left in the environment by decentralised actions stimulate the performance of subsequent actions by the same or a different agent. In a more applied sense, Klyubin *et al.* [28] treated agent’s sensors as extracting information and actuators as having the capability to “imprint” information on the environment, thus viewing agents as creating, maintaining and making use of various information “flows”. For example, the individuals within a swarm can put some information out into the environment, then retrieve it at a later point in time by sensing — i.e., individuals do not have to keep all of the information internally and can share a distributed collective memory through interactions with the environment or other individuals.

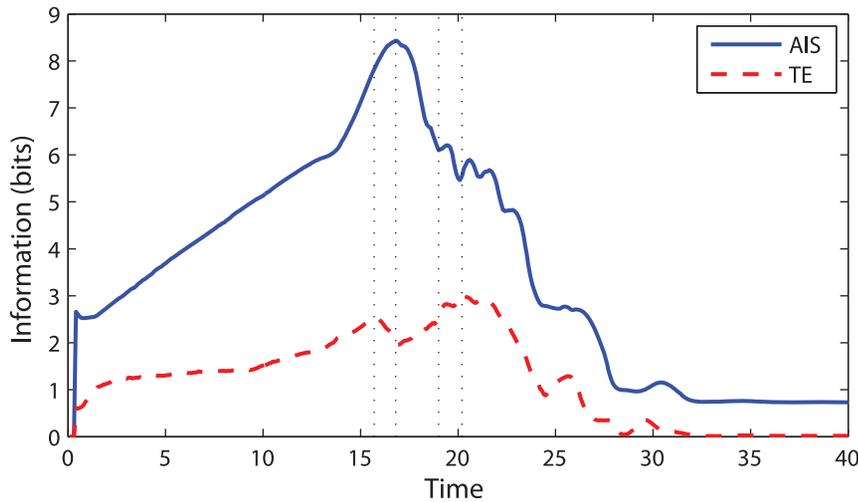
One may take a causation approach to measuring memory by computing causal information flows using interventionist approach of Ay and Polani [25]. In other words, one would attempt to impose on source variables and determine the changes in the destinations brought about by these impositions. For instance, if a swarm model is described by differential equations, one may estimate the effects of interactions between individuals by modifying terms of the model. In this work, however, we take a simpler approach to measuring memory via information storage, without causal flows.

To re-iterate, we hypothesise that AIS captures the active/predictive collective memory within the swarm while TE measures information cascades. To verify these hypotheses, we explore two scenarios. Our first experiment checks how different local initial perturbations affect a single swarm. The second experiment introduces a different type of perturbations, brought about by three separate but merging swarms. We use a Lagrangian model for modelling and simulating aggregations of discrete individuals. Each individual responds to its neighbours in three concentric zones with repulsion, orientation, or attraction, respectively [29–32]. The experiments quantitatively confirm our conjectures by tracing AIS and TE over time. The observed local and global maxima of these measures allow us to identify different elements of swarm dynamics (see Movie S1, S2, S3, S4 for the videos).

## Results

Initially, the individuals in the centre of the swarm are not affected by changes at the swarm’s periphery. As the changes propagate deeper, more and more individuals get engaged in collective computation and the collective memory grows, creating coordinated motion. When the majority of individuals are dynamically coordinated, average AIS of the swarm reaches its maximum (Fig. 1 at  $T=16.8$ ).

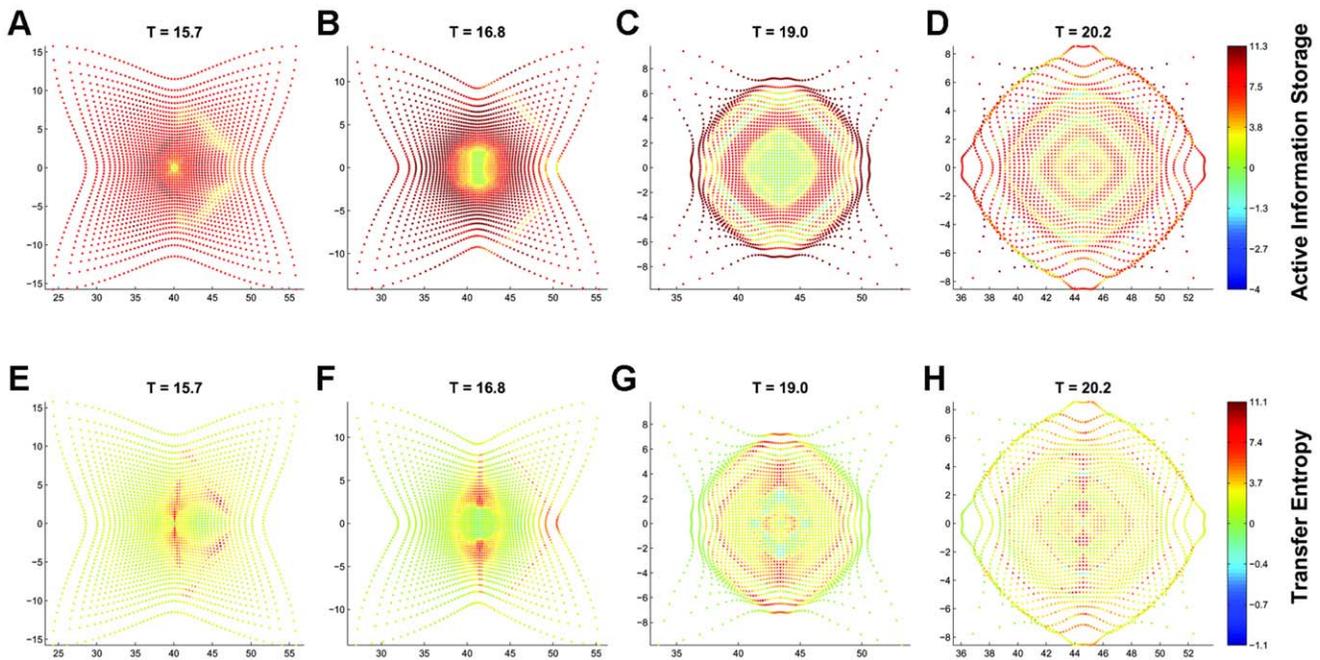
Figures 1, 2 and Movie S1 trace information dynamics over time, and show that local AIS can be positive and negative. Positive local AIS indicates that the past informs about the next state, while negative values indicate that the past misinforms about the next state [18,21]. Negative local storage means that an individual’s movement is unusually strongly influenced by other individuals (via high transfer) at this time, given the past history of that individual.



**Figure 1. Information storage and transfer over time for a swarm initially in a square configuration.** Shown here are the average active information storage (AIS) per particle and average transfer entropy (TE) per particle pair.  
doi:10.1371/journal.pone.0040084.g001

We argue that negative local AIS represents processing of new data propagated from elsewhere. For example, in Fig. 2C some of individuals in the centre are trying to compute their next state while being influenced by competing signals from their neighbours. In such situations, their past is misinformative about the next state. This results in the overall AIS decrease from the maximum. This decrease may be interrupted when the misinformed individuals are outnumbered by the individuals moving coherently. However, as the swarm synchronises to a steady motion, AIS (i.e. active collective

memory) converges to a positive constant near zero. This can be seen by writing the joint probability  $p(x_{n+1}, x_n^{(k)})$  as  $p(x_{n+1}|x_n^{(k)})p(x_n^{(k)})$  according to Bayes' Rule, making the log term in Equation 6 equal to  $\log_2 \frac{p(x_{n+1}|x_n^{(k)})p(x_n^{(k)})}{p(x_n^{(k)})p(x_{n+1})} = \log_2 \frac{p(x_{n+1}|x_n^{(k)})}{p(x_{n+1})}$ . When the swarm is in steady motion,  $p(x_{n+1}|x_n^{(k)}) \approx p(x_{n+1})$ , making the log term approximately zero. The 'bell' shaped curve of AIS is reminiscent of many complexity curves [33,34], indicating



**Figure 2. Local information dynamics for a swarm initially in a square configuration.** (Top) Local average information storage through the swarm at key time steps. (Bottom) Local average information transfer in a swarm at key time steps. The individuals' two-dimensional positions are plotted on the  $x$  and  $y$  axes, and their colours represent the AIS and TE values in bits, as shown by the scale on the colour bar. Note that the axes scales are adjusted, and the colours are scaled differently for positive and negative values.  
doi:10.1371/journal.pone.0040084.g002

that the most complex collective behaviour can be characterised information-theoretically.

Now we turn our attention to the communication aspect of computation, modelling information cascades by TE. As the swarm begins to ‘absorb’ the initial changes originated at the periphery, the first wave results in slightly increasing overall TE ( $T < 10.0$ ). The moment the wave reaches the centre, some individuals there acquire high local TE, being strongly influenced by their neighbours. Their new dynamics generates a new information wave spreading outwards through the swarm, achieving a local maximum at  $T = 15.7$  followed by a local minimum when it dissipates at  $T = 16.8$ . At this time AIS attained its global maximum, and the computation is non-trivial involving both memory and communication.

Local TE can also be positive or negative [20,21] (Fig. 2, bottom row, and Movie S2). Positive local TE means that the source agent is informative about the next state of the destination, given the destination’s history (the movement is strongly affected by its neighbours). Negative TE indicates that the source misleads an observer (when the individual is either exhibiting strong independent motion or is under the collective influence of several neighbours rather than the coherent influence of a single neighbour, e.g.  $T = 19.0$ ). These information dynamics suggest that transfer alternates with storage. Indeed, Figure 2 shows in most cases, areas of high local storage often have low or negative local transfer and vice versa.

Individuals that begin to move coherently (i.e. have comparably high local TE) form a front of a cascade as seen at  $T = 20.2$ . At that time, TE reaches its global maximum because the formed cascades dominate incoherent individuals. Not surprisingly, this stage has followed the time when memory (AIS) was highest. Eventually, the cascades help to coordinate the swarm, creating a steady configuration. TE decreases as the swarm ‘crystallises’.

Our second experiment (Fig. 3 & 4, and Movie S3 & Movie S4) models three swarms that eventually start interacting with each other. Thus, this experiment allows us to model different boundary perturbations.

When the swarms start to interact and the individuals become more dynamically coordinated, the overall local storage increases until it reaches global maximum at  $T = 76.0$  (Figure 4A–C). This is the moment when all three swarms merge into a single

coordinated entity, confined within a spatial extent that will not change significantly past this point. Importantly, the majority of individuals are dynamically coordinated at this instant, which is followed by several ‘waves’ (Figure 4D–H) that interleave groups of individuals with higher and lower local AIS. This process non-monotonically reduces AIS, while the swarm converges to the state of synchrony, with AIS being near zero (similar to the first experiment).

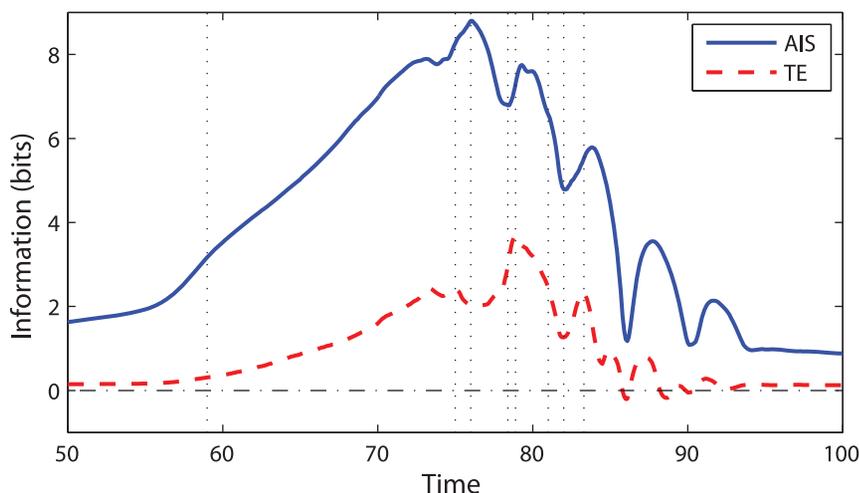
Similar to the first experiment, as the swarms merge ( $T \leq 78.9$ ), the overall TE is mostly increasing because there are several significant information cascades ‘rippling’ through the swarm. This shows that as the swarm is merging into one group, the specific individuals are under the most influence from their neighbours. The maximum TE lags behind that of AIS (analogously to the first experiment) again highlighting rich computation in terms of both collective communication and memory.

Tracing TE reveals some new features. Firstly, when the swarms merge, while moving from left to right, we can observe asymmetry in local values: a front of negative local TE on the left (where individuals were ‘surprised’ because the direction of dominant attraction was roughly opposite to the current direction of the overall swarm motion), and a front of positive local TE on the right (where these two directions concurred).

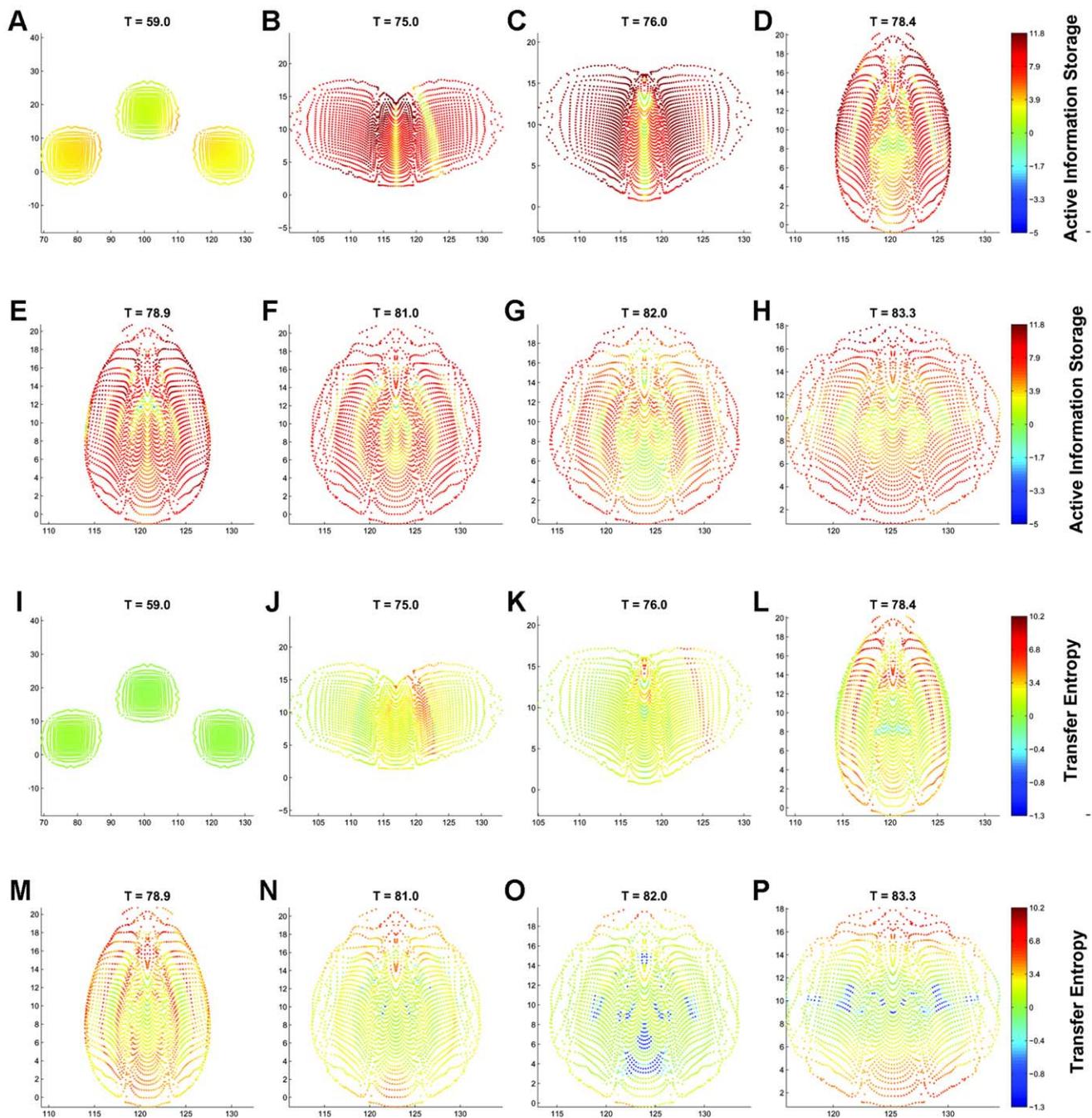
The second new feature is propagation of an initial asymmetry in swarms. For example, negative local TE at  $T = 82.0$  (Fig. 4O) are particularly visible. As the swarm progresses towards a steady state, there are further local fluctuations reducing the overall information storage and transfer values, showing that the distributed computation declines.

## Discussion

The reported results provide the first quantitative evidence (verified in a simulation setting) with a direct measure of information for the long-held conjecture that the information cascades occur in waves rippling through the swarm. The cascades can be observed via coherent changes in local TE, and are akin to information cascades in other systems, e.g. gliders in cellular automata [20]. Our characterisation deals with weak and ambiguous external stimuli by incorporating both positive and



**Figure 3. Information storage and transfer over time for a swarm initially consists of three squares in a checker configuration.** Shown here are the average active information storage (AIS) per particle and average transfer entropy (TE) per particle pair. doi:10.1371/journal.pone.0040084.g003



**Figure 4. Local average information dynamics in a swarm at key time steps.** A-H, active information storage, I-P, transfer entropy. The swarm initially consists of three squares in a checker configuration. The individuals' two-dimensional positions are plotted on the  $x$  and  $y$  axes, and their colours represent the AIS and TE values in bits, as shown by the scale on the colour bar. Note that the axes scales are adjusted, and the colours are scaled differently for positive and negative values.  
doi:10.1371/journal.pone.0040084.g004

negative local TE. In contrast to previous studies, information cascades are not just observed as changes in behaviours and activities, but are rather rigorously determined and computed.

In addition, we introduced a novel information-theoretic characterisation of swarm's collective memory, which is identified with AIS. Higher values of AIS are associated with higher levels of dynamic coordination. This study reveals different ways in which a particle's spatial position is dynamically related to its information processing role.

Collective communication and memory are two necessary elements of distributed computation (in addition to information modification [21,35]). The information-theoretic approach clearly separates different elements of distributed computation taking place in swarms, filtering and predicting important hot spots (e.g. a cascade's wavefront, collective memory's core, etc.). In addition, this framework may reveal new biological/social principles that govern coherent aggregation of living organisms (e.g. maximal

information transfer tends to follow the stage with maximal collective memory).

## Methods

We use a three-zone swarming model that features continuous, concentric circular and overlapping zones with smooth transitions. In an appropriate limit corresponding to a swarm consisting of a large number of individuals, the dynamics of the system is governed by a system of partial differential equations describing the density and velocity of the swarm [36]. To perform simulations, the density and velocity fields are systematically discretized into individuals with two-dimensional position vector  $\vec{s}_i$ , velocity  $\vec{v}_i$  and acceleration  $(d/dt)\vec{v}_i$ . For this model, individuals turn toward a desired direction,

$$\frac{d}{dt} \vec{v}_i = \kappa(\vec{v}_{d,i} - \vec{v}_i), \quad (1)$$

where  $\kappa$  is a turning rate parameter and  $\vec{v}_{d,i}$  is the desired direction of the  $i^{\text{th}}$  individual. The desired direction is a linear combination of influences,

$$\vec{v}_{d,i} = \vec{v}_{r,i} + \vec{v}_{o,i} + c_a \vec{v}_{a,i}, \quad (2)$$

where  $\vec{v}_{r,i}$ ,  $\vec{v}_{o,i}$  and  $\vec{v}_{a,i}$  are the influences from the zones of repulsion, orientation and attraction, respectively and are given by:

$$\vec{v}_{r,i} = \sum_{j=1}^N -\frac{1}{8\pi\sigma_1^4} \vec{s}_{ij} \exp(-|\vec{s}_{ij}|^2/4\sigma_1^2), \quad (3)$$

$$\vec{v}_{a,i} = \sum_{j=1}^N \frac{1}{64\pi\sigma_3^6} \vec{s}_{ij} |\vec{s}_{ij}|^2 \exp(-|\vec{s}_{ij}|^2/4\sigma_3^2), \quad (4)$$

and

$$\vec{v}_{o,i} = \frac{\sum_{j=1}^N \frac{1}{4\pi\sigma_2^2} \exp(-|\vec{s}_{ij}|^2/4\sigma_2^2) \vec{v}_j}{\sum_{j=1}^N \frac{1}{4\pi\sigma_2^2} \exp(-|\vec{s}_{ij}|^2/4\sigma_2^2)} \quad (5)$$

where  $\vec{s}_{ij} := \vec{s}_j - \vec{s}_i$ , the relative position of individuals  $i$  and  $j$ . The lengths  $\sigma_1 < \sigma_2 < \sigma_3$  represent the sizes of the repulsion, orientation and attraction zones, respectively. The parameter  $c_a$  in Equation 2 controls the relative importance of attraction over repulsion. An individual will be influenced to move away from other individuals that are within the innermost zone of repulsion, to align with individuals in the central zone of orientation and to move toward individuals in the outer zone of attraction. The constant  $c_a$  specifies the importance of attraction relative to orientation and repulsion.

We integrated the individual trajectories using the `scipy.integrate.odeint` Python package so that they are numerically resolved to a relative error of  $10^{-6}$ . To construct time series, we subsampled the trajectories at time intervals of  $\delta\tau$ . We start our investigation with individuals in a square configuration  $49 \times 49$  in size, or individuals in three squares of checker configuration initially, each square is  $28 \times 28$  in size.

AIS for agent  $X$  is the local mutual information from its semi-infinite past  $x_n^{(k)} = \{x_n, x_{n-1}, \dots, x_{n-k+1}\}$  (as  $k \rightarrow \infty$ ) to its next state  $x_{n+1}$  at time step  $n+1$  [18]:

$$a_X(n+1) = \lim_{k \rightarrow \infty} \log_2 \frac{p(x_n^{(k)}, x_{n+1})}{p(x_n^{(k)})p(x_{n+1})}, \quad (6)$$

with  $a_X(n,k)$  representing an approximation with finite history length  $k$ . The overall AIS is the average  $A_X(k) = \langle a_X(n,k) \rangle$ .

The local TE [20] from a source agent  $Y$  to a destination agent  $X$  is the local mutual information between the previous state of the source  $y_n$  and the next state of the destination  $x_{n+1}$ , *conditioned* on the past of the destination  $x_n^{(k)}$ . In this study, we also condition it on another contributor  $W$  to form the *conditional transfer entropy* [21]:

$$t_{Y \rightarrow X|W}(n+1) = \lim_{k \rightarrow \infty} \log_2 \frac{p(x_{n+1}|x_n^{(k)}, w_n, y_n)}{p(x_{n+1}|x_n^{(k)}, w_n)}. \quad (7)$$

Again,  $t_{Y \rightarrow X|W}(n,k)$  represents finite- $k$  approximation, and the overall TE is the average:  $T_{Y \rightarrow X|W}(k) = \langle t_{Y \rightarrow X|W}(n,k) \rangle$ .

To apply information dynamics to swarms, we accumulated the observations across agents and measured the state transitions with relative variables [37]. For local AIS, the variables in Eq. 6 are:  $x_n^{(k)} = \{\vec{v}_p^n - \vec{v}_p^{n-1}, |v|^n\}$ , and  $x_{n+1} = \{\vec{v}_p^{n+1} - \vec{v}_p^n, |v|^{n+1}\}$ . For TE, we do not take into account the speed in  $x_{n+1}$ , and  $y_n$  is the relative positions and velocities between two individuals, thus,  $y_n = \{\vec{s}_p^n - \vec{s}_p^{n-1}, \vec{v}_p^n - \vec{v}_p^{n-1}\}$ ,  $w_n = |v|^n$ ,  $x_n^{(k)} = \vec{v}_p^n - \vec{v}_p^{n-1}$ , and  $x_{n+1} = \vec{v}_p^{n+1} - \vec{v}_p^n$ .

For each individual, we compute local TE from all neighbours within a certain radius and average these values into the local TE for that individual. While each TE could be viewed as akin to a vector, i.e. having magnitude and relative direction from the source to the destination, these components are quite separate and the total information transferred to an individual does not add together in a simple vector-wise fashion. (Indeed, it adds together properly by considering incrementally conditioned transfer entropy terms, see [21]). As such, it is more meaningful to consider the average information received by the individual from each source. The average TE for the swarm is the average of these averages. For example, Figs. 2E–H show the local TE values for individuals at different times, while Fig. 1 traces the swarm average over time.

## Supporting Information

**Movie S1 Local active information storage (AIS) of swarm individuals over time for experiment 1.** We start the simulation with the swarm individuals in a square configuration  $49 \times 49$  in size. The top panel shows the average AIS per individual as depicted in Figure 2 in the main text, with the vertical line highlighting the specified time. The bottom panel shows the swarm dynamics at the specified time, the individuals' two-dimensional positions are plotted on the  $x$  and  $y$  axes. The local AIS values for each individual is shown here in different colours, according to the scale on the right. Note the colours are scaled differently for positive and negative values. (MOV)

**Movie S2 Local transfer entropy (TE) of swarm individuals over time for experiment 1.** We start the simulation with the swarm individuals in a square configuration  $49 \times 49$  in size. The top panel shows the average TE per particle pair as depicted in Figure 0 in the main text, with the vertical line

highlighting the specified time. The bottom panel shows the swarm dynamics at the specified time, the individuals' two-dimensional positions are plotted on the  $x$  and  $y$  axes. The local TE values for each individual is shown here in different colours, according to the scale on the right. Note the colours are scaled differently for positive and negative values.

(MOV)

**Movie S3 Local active information storage (AIS) of swarm individuals over time for experiment 2.** We start the simulation with the swarm individuals in three squares of checker configuration; each square is  $28 \times 28$  in size. The top panel shows the overall AIS as depicted in Figure 2 in the main text, with the vertical line highlighting the specified time. The bottom panel shows the swarm dynamics at the specified time, the individuals' two-dimensional positions are plotted on the  $x$  and  $y$  axes, with the colours of each individual denoting the value of its local AIS. The local AIS values for each individual is shown here in different colours, according to the scale on the right. Note the colours are scaled differently for positive and negative values.

(MOV)

**Movie S4 Local transfer entropy (TE) of swarm individuals over time for experiment 2.** We start the simulation with the swarm individuals in three squares of checker

configuration; each square is  $28 \times 28$  in size. The top panel shows the overall TE as depicted in Figure 4 in the main text, with the vertical line highlighting the specified time. The bottom panel shows the swarm dynamics at the specified time, the individuals' two-dimensional positions are plotted on the  $x$  and  $y$  axes, with the colours of each individual denoting the value of its local AIS. The local AIS values for each individual is shown here in different colours, according to the scale on the right. Note the colours are scaled differently for positive and negative values.

(MOV)

## Acknowledgments

We would like to thank the High Performance Computing and Communications Centre (<http://www.hpccc.gov.au/>) for the use of their supercomputer clusters in performing the experiments for this paper. Part of this work was completed during a stay of JMM at the CSIRO in Sydney, Australia, and JMM thanks CSIRO for support and hospitality.

## Author Contributions

Performed the experiments: XRW JTL JMM LFR. Analyzed the data: XRW MP. Contributed reagents/materials/analysis tools: XRW JTL JMM LFR. Wrote the paper: XRW MP. Conceived the experiments: XRW JTL LFR. Designed the experiments: XRW JTL JMM.

## References

- Lissaman PBS, Shollenberger CA (1970) Formation flight of birds. *Science* 168: 1003–1005.
- Parrish JK, Edelstein-Keshet L (1999) Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* 284: 99–101.
- Sinclair ARE, Norton-Griffiths M, editors (1979) *Serengeti: Dynamics of an Ecosystem*. University of Chicago Press.
- Camazine S, Deneubourg JL, Franks NR, Sneyd J, Theraulaz G, et al. (2003) *Self-Organization in Biological Systems*. Princeton University Press.
- Giraldeau LA, Valone TJ, Templeton JJ (2002) Potential disadvantages of using socially acquired information. *Proc R Soc B* 357: 15591566.
- Potts WK (1984) The chorus-line hypothesis of manoeuvre coordination in avian flocks. *Nature* 309: 344–345.
- Couzin ID, James R, Croft DP, Krause J (2006) Social organization and information transfer in schooling fishes. In: Brown C, Laland K, Krause J, editors, *Fish Cognition and Behavior*, Blackwell Publishing, Fish and Aquatic Resources. 166–185.
- Radakov DV (1973) *Schooling in the ecology of fish*. New York: John Wiley & Sons Inc. Translated from Russian by H. Mills.
- Kastberger G, Schmelzer E, Kranner I (2008) Social waves in giant honeybees repel hornets. *PLoS ONE* 3: e3141.
- Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashions, custom, and cultural change as informational cascades. *J Polit Econ* 100: 992–1026.
- Galef BG Jr, Giraldeau LA (2001) Social influences on foraging in vertebrates: causal mechanisms and adaptive functions. *Anim Behav* 61: 3–15.
- Dall SRX, Giraldeau LA, Olsson O, McNamara JM, Stephens DW (2005) Information and its use by animals in evolutionary ecology. *Trends Ecol Evol* 20: 187–193.
- Schreiber T (2000) Measuring information transfer. *Phys Rev Lett* 85: 461–464.
- Katz Y, Tunstrom K, Ioannou CC, Huepe C, Couzin ID (2011) Inferring the structure and dynamics of interactions in schooling fish. *Proc Natl Acad Sci USA* 108: 18720–18725.
- Couzin ID (2007) Collective minds. *Nature* 445: 715–715.
- Polani D (2009) Information: currency of life? *HFSP J* 3: 307–316.
- Sperati V, Trianni V, Nolfi S (2008) Evolving coordinated group behaviour through maximization of mean mutual information. *Swarm Intelligence* 2: 73–95.
- Lizier JT, Prokopenko M, Zomaya AY (2012) Local measures of information storage in complex distributed computation. *Information Sciences* 208: 39–54.
- Lizier JT, Prokopenko M, Zomaya AY (2007) Detecting non-trivial computation in complex dynamics. In: Almeida e Costa F, Rocha LM, Costa E, Harvey I, Coutinho A, editors, *Proceedings of the 9th European Conference on Artificial Life (ECAL 2007)*, Lisbon, Portugal. Berlin/Heidelberg: Springer, volume 4648 of *Lecture Notes in Artificial Intelligence*, 895–904.
- Lizier JT, Prokopenko M, Zomaya AY (2008) Local information transfer as a spatiotemporal filter for complex systems. *Phys Rev E* 77: 026110.
- Lizier JT, Prokopenko M, Zomaya AY (2010) Information modification and particle collision in distributed computation. *Chaos* 20: 037109.
- Lizier JT, Heinzele J, Horstmann A, Haynes JD, Prokopenko M (2010) Multivariate informationtheoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *Journal of Computational Neuroscience* 30: 85–107.
- Wibral M, Rahm B, Rieder M, Lindner M, Vicente R, et al. (2011) Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks. *Progress in Biophysics and Molecular Biology* 105: 80–97.
- Chicharro D, Ledberg A (2012) When two become one: The limits of causality analysis of brain dynamics. *PLoS ONE* 7: e32466.
- Ay N, Polani D (2008) Information flows in causal networks. *Advances in Complex Systems* 11: 17–41.
- Lizier JT, Prokopenko M (2010) Differentiating information transfer and causal effect. *European Physical Journal B* 73: 605–615.
- Grasse PP (1959) La reconstruction du nid et les coordinations interindividuelles chez *Bellicositermes natalensis* et *Cubitermes* sp. La theorie de la stigmergie: Essai d'interpretation des termites constructeurs. *Insectes Sociaux* 6: 41–83.
- Klyubin AS, Polani D, Nehaniv CL (2004) Tracking information flow through the environment: Simple cases of stigmergy. In: Pollack J, Bedau M, Husbands P, Ikegami T, Watson RA, editors, *Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*. The MIT Press, 563–568. doi:10.2307/40041279.
- Couzin ID, Krause J, James R, Ruxton GD, Franks N (2002) Collective memory and spatial sorting in animal groups. *J Theor Biol* 218: 1–11.
- Huth A, Wissel C (1992) The simulation of the movement of fish schools. *J Theor Biol* 156: 365–385.
- Lukeman R, Li YX, Edelstein-Keshet L (2010) Inferring individual rules from collective behavior. *Proc Natl Acad Sci USA* 107: 12576–12580.
- Vicsek T, Czirok A, Ben-Jacob E, Cohen I, Shochet O (1995) Novel Type of Phase Transition in a System of Self-Driven Particles. *Phys Rev Lett* 75: 1226–1229.
- Lizier JT, Pritam S, Prokopenko M (2011) Information dynamics in small-world Boolean networks. *Artificial Life* 17: 293–314.
- Prokopenko M, Boschetti F, Ryan AJ (2009) An information-theoretic primer on complexity, selforganization, and emergence. *Complexity* 15: 11–28.
- Langton CG (1990) Computation at the edge of chaos: phase transitions and emergent computation. *Physica D* 42: 12–37.
- Miller JM, Kolpas A, JuchemNeto JP, Rossi LF (2011) A continuum three-zone model for swarms. *B Math Biol* 1: 1–26.
- Wang XR, Miller JM, Lizier JT, Prokopenko M, Rossi LF (2011) Measuring information storage and transfer in swarms. In: Lenaerts T, Giacobini M, Bersini H, Bourgine P, Dorigo M, et al., editors, *Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems (ECAL 2011)*. Paris, France: MIT Press, 838–845.