15TH ICCRTS

"The Evolution of C2"

Phase synchronization of an ensemble of weakly coupled oscillators: A paradigm of sensor fusion.

Topics: (6) Modelling and Simulation; (1) Concepts, Theory, and Policy

Ariën J. van der Wal

POC: Ariën J. van der Wal

Netherlands Defence Academy (NLDA)
P.O. Box 10000, 1780 CA Den Helder, The Netherlands
Tel: +31 223657363 / Fax: +31 76 527 3259
AJ.vd.Wal@nlda.nl

PHASE SYNCHRONIZATION OF AN ENSEMBLE OF WEAKLY COUPLED OSCILLATORS: A PARADIGM OF SENSOR FUSION

Ariën J. van der Wal

Netherlands Defence Academy, P.O. Box 10000, 1780 CA Den Helder, The Netherlands
e-mail: AJ.vd.Wal@nlda.nl

ABSTRACT

One of the most challenging phenomena that can be observed in an ensemble of interacting agents is that of self-organisation, *viz.* emergent, collective behaviour, also known as synergy. The concept of synergy is well-known in the artificial intelligence community, in social and management science, and also in a the military C2 community, *e.g.* in describing sensor fusion. The idea often loosely phrased as '1+1>2', strongly suggests that it is possible to make up an ensemble of similar agents, assume some kind of interaction and that in such a system 'synergy' will automatically evolve. In a more rigorous approach the paradigm may be expressed by identifying an ensemble performance measure that yields more than the sum of the individual performance measures of the constituents.

The aim of the present study is to discuss in a simple conceptual model system under what circumstances self-organization is feasible and to discuss what type of agents and interactions are minimally required to induce synergy among agents. As a case in point we discuss the emergent phase coherence of a multi-oscillator system with non-linear all-to-all coupling between the oscillators. In the thermodynamic limit this system shows spontaneous organization. Simulations indicate that also for finite populations that are not completely connected phase synchronization spontaneously emerges if the interaction strength is strong enough.

1. INTRODUCTION

Sensor fusion has been studied over the past two decades. Although it is clear that many identical sensors in cleverly arranged set-ups may increase resolution and are able to track multiple targets, and dissimilar sensors may be able to refine each others observations by complementing each others capabilities, from the information science point of view the added value of merging the results of many sensors, connected in networks has not been as successful as expected. One of the reasons for this is the absence of a clear view on what can exactly be expected in terms of a better assessment of the environment by combining the information of multiple similar or dissimilar sensors. Because the metric of sensor performance is lacking, it is difficult to demonstrate its real virtue. In addition the dimensionality (space, time, features) of the fusion problem makes it difficult to define a universal description of the process.

From the theoretical side there have been several attempts to define measures of synergy, *viz.* probabilistic and fuzzy belief measures that express the added value ("synergy") of cooperating ensembles of many sensors. It is not obvious at all that by just adding more sensors to a network or having more than one sensor observation of the same event is automatically improving situational awareness. The management perception of synergy ("the synthesis of partial results is more than their sum") is in practice hard to demonstrate.

The elementary question therefore is: Can we think of a mechanism or paradigm that shows in a logical way the added value of synergy. We could even think of a more abstract way of posing this question: Is it possible to demonstrate in a realistic physical model that by combining simple, elementary "agents", we may create new performance measures of the resulting ensemble, *i.e.* new properties pertaining to the ensemble as a whole that are not present in the individual agents ?

In the present article we will discuss a candidate physical model that is able to demonstrate the emergence of new characteristics of an ensemble that are not present in its separate constituents. The occurrence of spontaneous synchronization of an ensemble of phase oscillators with non-linear interaction is a paradigm that demonstrates added value in terms of degree of synchronization and robustness if the interaction between the oscillators is strong enough. These properties are not defined for the individual phase oscillators.

Finally we note that the theoretical model is in fact also in practice important, because synchronization is the key concept necessary for distribution of time in a network of sensors and thus, on the physical layer of the network suggests a way to order time-stamped events and distribute time, with the need of neither a central fusion centre nor complicated multiple access or routing techniques, or an elaborate time distribution protocol. This in turn makes robust synchronization and communication in large, scaleable networks based on analog communication feasible with a time granularity of the order of one microsecond or less.

2. MODEL

2.1 Network topology

In real sensor networks, and especially in wireless sensor networks, we have a number of contradicting, competing operational requirements. In the following we will primarily consider networks that are deployed by the military with the aim to survey a certain area, *e.g.* to perform a detection function: sentry nodes. The sensor nodes are dropped in the area with a certain average density and will have to operate autonomously: power backup is contained in the sensor and in principle it is possible to re-supply energy by solar cells.
The basic network topology is fixed although movement in space cannot be precluded due to external, environmental changes. However, in the present model the nodes are not moving around autonomously. Exchange of information between nodes is assumed to be by wireless r.f. communication, although the principles outlined here extend to IR or even acoustic communication as well, with obvious adjustments for bandwidth, delays, throughput, energy consumption, *etc*.

Each sensor has very limited resources in terms of energy consumption and power, so that the network topology cannot be rigorously maintained for all times, and modelling of the description of the network requires switching topologies, resulting in a geometric random graph description. In this way the overall lifetime of the distributed sensor network can be significantly extended. At the same time a geometric random graph approach introduces robustness in the network modelling: nodes may malfunction, die or revive without compromising the overall operation of the network as a whole.
Sensors deployed in large-scale networks must by definition be extremely cost effective: they are only used once, should be relatively simple in terms of hardware and computational capabilities, yet should be accurate and dependable as detectors, and be able to communicate their results over long observation periods.
Because of the limited computational capabilities of the single sensor node, time distribution and the fusion of data has to be done by specialized, central data fusion and synchronization nodes. In large networks the collection and processing of experimental data from sensor nodes and the diffusion of timing information to the sensors, if time stamping is required poses a significant computational burden on the network in terms of computer power, latencies and bandwidth.

The limitations posed by the simple sensor nodes in terms of energy consumption introduce a number of problems: First the sensor cannot be supposed to be in permanent contact with the network at all times: it will hibernate from time to time in order not to exhaust its energy supply too quickly. Therefore such a network must be robust against (re)connecting and disconnecting of sensors. In the second place it is desirable to have a *scalable* network, *i.e.* the nodes should be homogeneous and no special nodes are necessary for the data processing and fusion. This is an important motivation to look for non-local, distributed fusion schemes.

There is another reason why it is important in practice to avoid central fusion in case the sensor network is used for event detection, *e.g*. in case of military interdiction or in the detection of hazardous events: Centralized communication networks often suffer from overload and long latencies at the very moments that they should be able to respond quickly in a reliable way. An example is the performance of the recently introduced communication system C2000 for the emergency services (police, fire brigade and ambulance services) in the Netherlands: If a serious accident occurs the system quickly overloads, because all responding emergency services move to the same area and communication performance deteriorates down to unacceptable levels. This is especially true for large-scale heterogeneous networks. In [1] it was shown that in a wireless network with one-to-one links, the transport capacity per sensor node is proportional to $1/\sqrt{N\log N}$ for large *N*, where *N* is the number of sources. This result applies to a situation with many sources and many sinks. In contrast, the data-centric nature of sensor networks was exploited in [2], where it was shown that if we are dealing with a symmetric function of the nodes, *i.e.* if we are not concerned with the allocation of sensors and measurement, the transport capability scales as $1/\log N$.

2.2. Agents as oscillators

In the present model we distinguish between functionality of the network, the network topology and finally the nature of the agents. Because we focus on arriving at a common decision by randomly distributed sensors, we will reduce the sensors to their bare essential functionality: how can an agent communicate its local decision to all other agents in the network and make the network arrive at a common decision on the basis of all local decisions in a energy-effective way? In particular we are interested in the added value of this communication and census process: Is it possible to show that the network arrives at a common decision that is the result of all contributing sensors and is "better" (in the sense of *e.g*. quality, confidence level, belief, robustness) than a mere superposition of single sensor decisions? For this reason we will not look into the sensing process itself, but merely assume that an observable can be expressed in a scalar value and that this value can be translated in a characteristic of the agent. Note that the restriction of one observable per sensor can be easily dropped by extending the model to a vector model, where the vector consists of m independent measurements, either of m physical sensors per node, or of a time sequence of m observations by one sensor. Each node (agent) will be modelled as a simple phase oscillator with an eigenfrequency that scales with the result of the measurement taken at the node. The oscillators are coupled to each other by means of a non-linear interaction that depends on their phase difference. Under suitable conditions this non-linear interaction gives rise to synchronization between the interacting oscillators.

2.3 Phase oscillators

First we review the behaviour of an ensemble of non-linear coupled phase oscillators and then apply this model to the distributed fusion of sensor data. Oscillators are in general non-linear devices in the sense that both amplitude and phase evolve in time and are described by first order non-linear differential equations (DE). Even purely linear harmonic oscillation with low total harmonic distortion is in practice achieved by *e.g.* a Wien-bridge oscillator that uses a non-linear feedback element, such as an incandescent lamp [3].
Although actual oscillators are non-linear systems of at least second order [4], in the present discussion it suffices to describe their dominant dynamics by the Liénard equation:

$$\frac{d^2x}{dt^2} + f(x)\frac{dx}{dt} + g(x) = 0 \qquad (1)$$

where *x(t)* is the output of the oscillator.
The function *f(x)* determines the shape and amplitude of the oscillation, whereas *g(x)* determines the oscillation frequency.
The Liénard equation [5-9] is a generalization of the famous Van der Pol equation:

$$\frac{d^2x}{dt^2} - \mu(1-x^2)\frac{dx}{dt} + x = 0 \qquad (2)$$

This second order differential equation describes a non-conservative oscillator with non-linear damping. A Liénard system has a unique and stable limit cycle solution if the Eq (1) satisfies the following conditions: *f(x)* is an even function, *g(x)* is an odd function, *f(x)* is continuous with $f(0) < 0$ and *g(x)* is Lipschitz, $xg(x) > 0$ for $\forall x \neq 0$ and $\lim_{x\to\pm\infty} F(x) \equiv \lim_{x\to\infty}\int_0^x f(\xi)d\xi \to \pm\infty$ and $F(x)$ has exactly one positive root a, with $F(x) < 0$ for $0 < x < a$ and $F(x)$ for $x > a$ is strictly increasing.

In the following we will use the concept of the phase oscillator: A phase oscillator is described only by the time dependence of its phase $\frac{d\vartheta}{dt} = \omega_0$. In this very simple description we consider just the time evolution of the phase θ and neglect the dependence of amplitude *vs.* time. Although this may seem a gross simplification, one can show that the theoretical concept is very useful and reduces the order of the DE describing the evolution of θ in time to first order.

Our discussion of the phase oscillator model is along the lines presented in [10]. We consider a network consisting of *N* nodes composed of a sensor and described by an autonomous dynamical system. Each sensor is coupled to its *(N-1)* neighbours, although this requirement may be relaxed to also include sparser coupled networks. Following [10] we assume that the sensor operates as a detector. Each sensor decides if a detection is made by setting its fundamental frequency $\omega_i = \Omega_1$; if no event is detected it sets $\omega_i = \Omega_0$.

The network of coupled oscillators thus works as a distributed detector. Each sensor node is represented by an oscillator that is described by the following equation:

$$\frac{d\vartheta_k(t)}{dt} = \omega_k + \frac{K}{c_k}\sum_{j=1}^{N} a_{kj}F(\vartheta_j(t) - \vartheta_k(t)) \quad (k = 1,\ldots,N) \qquad (3)$$

where $\vartheta_k(t)$ is the state of the k-th sensor, or alternatively, the phase angle of the k-th oscillator. The initial value $\vartheta_k(0)$ is taken as a random number in $[0, 2\pi]$; The function *F(x)* describes the non-linear interaction between pairs of oscillators and is supposed to be odd: $F(x) = -F(-x)$; the *N*N* matrix elements a_{ij} describe the coupling of the network: $a_{ij} = 1$ if i and j are connected and $a_{ij} = 0$ otherwise. We assume that the network is undirected, *i.e.* $a_{ij} = a_{ji}$. K is the mutual coupling constant and c_i weighs the influence of all other sensors on the i-th sensor. It can be shown to be a measure of confidence, or alternatively, the SNR of the i-th sensor. If we take $F(x) = \sin x$, and take $a_{ij} = a_{ji} = 1$ for all pairs (i,j), and $c_i = 1$ for all I, the model is equivalent to the Kuramoto model [11], which has been extensively studied. For an excellent review see [12] and references contained therein.

$$\frac{d\vartheta_k}{dt} = \omega_k + \frac{K}{N}\sum_{j=1}^{N}\sin(\vartheta_j - \vartheta_k) \quad (k = 1,\ldots,N) \qquad (4)$$

Finally it is worth noting that the same non-linear Kuramoto DE occurs in the classical theory describing the nonlinear first order analog phase locked loop (PLL) [13]. The description of the extremely non-linear capture behaviour of a PLL boils down to solving a set of two first-order nonlinear DE, one of which is of the same type as Eq. (4), after fast varying components are averaged out [14,15].

2.4 Emergent behaviour: synchronization

For very low values of *N, e.g. N*=2 the solution of the phase evolution is relatively easy. However the analysis of a general *N* oscillator system is much more complex and can only be done by simulation.

In the thermodynamic limit ($N \longrightarrow \infty$) we can use the so-called mean-field approximation, and analyze the behaviour of the system as if it were a physical ensemble of spins. From physics we know that the existence of a mean-field is indicative of a phase transition, *e.g*. thermodynamic transitions such as melting (solid⟶ liquid), normal ⟶ superconducting, paramagnetic ⟶ (anti)ferromagnetic order. In analogy to these physical phenomena it can be shown that the Kuramoto system also displays a phase transition [16] .

The basic derivation showing the existence of a phase transition in the Kuramoto model for $N \longrightarrow \infty$ and K large enough is straightforward. In analogy to the theory of phase transitions we start by defining a so-called order parameter r, the 'mean field', which characterizes the phase transition:

$$r(t)e^{i\Theta(t)} = \frac{1}{N}\sum_{j=1}^{N} e^{i\vartheta_j(t)} \qquad (5)$$

As can be seen $r(t)$ is a real number built up by the superposition of the contributions of all neighbouring nodes in the network. Generally the contributions to the sum Eq. (5) have arbitrary magnitudes and phases, so that they add up incoherently and therefore their contributions to $r(t)$ will be negligible. In case that for a fraction of the ensemble the oscillator frequencies $\dot{\vartheta}_j(t)$ converge to the same value, say ω^*, the superposition is *coherent* and $r(t)$ will tend to a constant value (of the order 1). Another way to characterize this situation is that all oscillators are synchronized, *i.e*. their phases $\vartheta_j(t)$ are in general different, but their phase *rates* (angular frequencies) $\dot{\vartheta}_j(t)$ are equal.

The oscillators in the ensemble are phase-locked. This can mathematically be formulated by putting $\vartheta_j(t) = \omega^* t + \vartheta_j(0)$. We then have for large *N:*

$$\frac{1}{N}\sum_{j=1}^{N} e^{i\vartheta_j(t)} \approx re^{i(\omega^* t + \Theta(0))} \qquad (6)$$

$$\frac{d\vartheta_k}{dt} = \omega_k + Kr\sin(\Theta - \vartheta_k) \qquad (7)$$

We can now transform Eq. (7) to a reference frame rotating with $\dot{\Theta} = \omega^*$ by setting $\psi_k = \vartheta_k - \Theta$:

$$\frac{d\psi_k}{dt} = \omega_k - \omega^* - Kr\sin(\psi_k) \qquad (8)$$

Eq. (8) has synchronous and asynchronous solutions depending on the strength of the interaction.

2.5 Connectivity graph of the network

The analysis in section 2.4 was restricted to all-to-all connections between oscillators. It can be extended to include other network topologies, provided that the interactions are *symmetric, i.e.* the network topology corresponds to an undirected graph. Oscillators that are entrained by the mean field will synchronize and oscillate with a common frequency ω^*.

The mean field approximation Eq. (5) has to be modified to include the connectivity of the network:

$$r_k(t)e^{j\Theta_k(t)} = \frac{1}{N}\sum_{j=1}^{N} a_{kj} e^{j\vartheta_j(t)} \qquad (9)$$

Note that the *global* mean field parameter $r(t)$ is now replaced by a *local* order parameter $r_k(t)$ and that the resulting general differential equation for the phase evolution is slightly more involved than Eq. (7):

$$\frac{d\vartheta_k}{dt} = \omega_k + \frac{K}{c_k} r_k \sin(\Theta_k - \vartheta_k) \qquad (10)$$

Again we change the frame of reference $\psi_k = \vartheta_k - \omega_k^* t$, where ω_k^* is now the synchronization frequency of the subgroup to which the k-th oscillator belongs.

$$\frac{d\psi_k}{dt} = \omega_k - \omega_k^* - \frac{K}{c_k} r_k \sin(\psi_k) \qquad (11)$$

The mathematical derivation of the condition for the interaction strength *K* is straightforward. Multiplying both sides of Eq. (10) with c_k and summing over all *N* yields:

$$\sum_{k=1}^{N} c_k \frac{d\psi_k(t)}{dt} = \sum_{k=1}^{N} c_k (\omega_k - \omega_k^*) \qquad (12)$$

where we have used that $a_{kj} = a_{jk}$ and the fact that *F(x)* = sin*x* is odd: $F(x) = -F(-x)$. Therefore it follows that if the ensemble synchronizes, the phase rates of all oscillators in the ensemble $\frac{d\psi_k(t)}{dt}$ will converge to the same value for t ⟶ ∞:

$$\omega_C^* = \frac{\sum_{k=1}^{N} c_k \omega_k}{\sum_{k=1}^{N} c_k} \qquad (13)$$

It is worth noting that the value ω_C^* pertains to all neighbouring oscillators within the same synchronous cluster *C*. More than one synchronous cluster could emerge in a general topology with different ω_C^*. In the limit of convergence each oscillator in such a synchronous cluster *C* satisfies:

$$\dot{\psi}_k = \omega_k - \omega_C^* - K r_k \sin(\vartheta_k - \omega_C^* t) \qquad (14)$$

from which follows the condition of synchronization

$$K r_k > \left| \omega_k - \omega_C^* \right| \qquad (15).$$

Eq. (15) implies that if the interaction strength *K* inside the cluster *C* is strong enough, intra-cluster synchronization will occur. If the network graph is completely connected, *i.e.* $a_{ij} = a_{ji} = 1$ for all pairs (i , j), *global* synchronization will occur. Although r at t = 0 is generally small because all oscillators initially each have their own frequency and are incoherent, more and more oscillators synchronize, *i.e.* oscillate with the common frequency ω^*. This partial synchronization entraines still more oscillators, further increasing r until almost all oscillators within *C* are coherent as reflected by r approaching the value of c / *N*, where c represents the size of cluster *C*. It can be shown that the equilibrium is stable for these values of K. In the case of global synchronization the value r can be seen as an order parameter in analogy to the magnetization in the paramagnetic ⟶ ferromagnetic phase transition in solid state physics.

On the other hand the network cannot synchronize if there exists an oscillator for which $Kd < \left| \omega_i - \omega^* \right|$ since $r \le d$, where d represents the degree of the network. This implies that the larger d, the easier the network will synchronize with a given interaction strength *K* and that in a fully connected network already relatively weak interaction will suffice to achieve synchronization among the *N* oscillators. By numerically solving the set of differential equations Eq. (10) for finite *N* and for *K* large enough global synchronization is observed for networks corresponding to *regular* graphs of different degrees d. The value of the order parameters $r_k(t)$ quickly converges to a value slightly below d/*N*. The rate of convergence increased with increasing d.

2.6 Sensors as detectors

The previous analysis can now be exploited to use the sensor network as a distributed detector: In analogy to the analysis of Ref. [10] we define two alternatives for each sensor: $\omega_i = \Omega_1$ if a sensor detects an event and $\omega_i = \Omega_0$ if no event is detected. This analysis can be extended to the simultaneous detection of more than one type of event. If we view detection as a stochastic process with a probability of detection p_0, we may conclude from the previous discussion that if a sensor network is sufficiently large and if the phase rate converges to ω^*, then the value of ω^* is given by the expectation value of ω_i:

$$\frac{d\vartheta^*}{dt} = \omega^* = \frac{\sum_{k=1}^{N} c_k \omega_k}{\sum_{k=1}^{N} c_k} \equiv p_0 \Omega_1 + (1-p_0)\Omega_0 \qquad (16)$$

The evolution in time of the phases of the N oscillators is then given by

$$\vartheta_j(t) = \omega^* t + \vartheta_j(0) \qquad (17),$$

where $\vartheta_j(0)$ the initial phase of each oscillator is given by the alternative

$$\vartheta_j(0) = \begin{cases} \Theta_0 = \arcsin\left(\dfrac{\omega^* - \Omega_0}{Kr}\right) & with \quad probability = 1 - p_0 \\ \Theta_1 = \arcsin\left(\dfrac{\omega^* - \Omega_1}{Kr}\right) & with \quad probability = p_0 \end{cases} \qquad (18)$$

This statistical approach gives rise to a Bernoulli distribution for r and from this the expectation value of r can be calculated by imposing that the r calculated in the mean field approximation Eq.(6) coincides with the expectation value. The solution of this implicit equation yields values of r in close agreement with $r \approx d$, the network degree.

In Eq. (13) the common oscillation frequency ω_C^* of the synchronized cluster C is expressed as a weighted average over the eigenfrequencies of the participating oscillators. It seems therefore natural to identify the coefficients c_k as the level of confidence one associates with of the k-th sensor. In a statistical sense we can therefore write $c_k = 1/\sigma_k^2$, where σ_k^2 is the variance of a single measurement. In this way the various repetition rates of sensors could be taken into account, which is especially important when dealing with a wide range of timescales. By identifying the weights c_k as *belief* or *confidence* measures sensors that produce more accurate results contribute more to the common synchronization frequency and thus to the sensor fusion process than sensors that are less accurate.

Finally we discuss a way to extend the use of multiple sensors beyond detection, *e.g.* for more accurately measuring a scalar field of sensor values. The sensor modelling discussed so far has exploited the global synchronization as a means to arrive at consensus among the N individual sensors in the network. If one considers the detection of intruders in the area where the sensor network is deployed this may be desirable. However in the case one wants to do a measurement of a quantity with this network and one is interested in learning how a certain scalar field varies over the surface covered by the sensors, one has to be more subtle. Although an exact analysis is outside the scope of the present work we suggest to modify the Kuramoto model Eq.(4) in such a way that the coupling between each sensor pair is weighted with a factor $\exp\left[\dfrac{-(\vartheta_j - \vartheta_k)^2}{2(\Delta\mu)^2}\right]$, *i.e.* that the function $F(x)$ in Eq.(3) is taken as $F(x) = e^{-x^2/2(\Delta\mu)^2} \sin x$ instead of the Kuramoto term $F(x) = \sin x$. The inclusion of this extra weight ensures that clusters are formed in space, indicative of similar measurements. In this case global synchronization is impeded and local clusters are formed. A similar local clustering can be exploited by collectively varying the transmission power at specific time intervals. With these methods much more information pertaining to the scalar field can be extracted. The idea is similar to the morphogenetic neuron [17], where instead of linear superposition of scalar outputs of neighbouring neurons in a Hopfield neural net instead *non-linear basis functions* are combined in a synapse, before the threshold function is applied and the output is generated. In this way it has been shown that more universal data geometries could be mapped by the neural network. Future research is necessary to dynamically vary the phase mismatch parameter $\Delta\mu$ and in this way obtain a better spatial resolution of the detection, or the measured scalar field, without the need for central fusion or extensive communication overhead.

3. SYNERGY: THE ORDER PARAMETER OF PHASE SYNCHRONIZATION

Synergy is commonly defined as the effect that if agents work together, the result of this co-operation is greater than the sum of the results produced by the individual agents. Although this is an interesting abstract theoretical concept, in practice it is not so obvious that this expected effect really exists.

Sensor fusion is a process that is aimed at improving the quality of the observation process by combining the data streams originating from different sensors. Although is in practice not easy at all, in an abstract way it is easily conceived that combining the outputs of many sensors will increase our understanding of the observable world around us. This has always been the drive to study sensor fusion. However apart from the fact that multiple sensors of the same type provide better statistics and thus more accurate results and different complementary sensors provide more, independent information and therefore potentially a better awareness, it is by no means obvious that adding more sensors automatically increases our understanding of the world around us, in other words: that there is in fact synergy in sensor fusion .

In an abstract way the concept of synergy may be formulated in several ways, *e.g.*

$$g(A \cup B) \geq \max(g(A), g(B)) \tag{19}$$

or

$$g(A \cup B) \geq g(A) + g(B) \tag{20}$$

From this formulation it is clear that in order to model synergy, we need non-linear operators.

The earliest attempts to combine measurements from multiple sources in a non-linear way are by Bayes [18]. He introduced the notion of conditional probability *Prob* (A|B), defined by:

$Prob(A|B) = \dfrac{Prob(A \cap B)}{Prob(B)}$ *i.e.* the probability of A, given that event B has occurred. This definition is easily extended to n observations obtained by n sensors. There are a number of difficulties connected with the application of the Bayesian sensor fusion formula:

- difficulty of assigning *a priori* probabilities
- complexity when there are multiple hypotheses and/ or multiple conditional events
- requirement that hypotheses have to be exhaustive and mutually exclusive
- absence of uncertainty modelling

In trying to find an appropriate way to model fusion and take advantage of the nonlinearity of the process, Dempster and Shafer (DS) [19,20] created a generalization of Bayesian theory that allows for the incorporation of uncertainty by using (overlapping) probability *intervals* and uncertainty modelling to determine the likelihood of hypotheses based on multiple evidence [21]. The essential generalization of DS theory is that not *all* hypotheses need to be mutually exclusive as in the Bayesian theory. In DS fusion evidence is assigned both to single and more general propositions, instead of assigning a probability to each hypothesis like in Bayesian theory.

Noting that belief and plausibility measures are both examples of Sugeno's [22] λ-fuzzy measure g_λ, the question arises whether it is possible to combine the intuitive ideas on sensor fusion and the properties of g_λ. We will show that in contrast the basic probability assignment in DS theory, fuzzy g_λ measures can indeed be utilized for the problem under consideration. We will take a closer look at this in the following and propose to view the multi-sensor fusion process in terms of a synergy between (sets of) sensors that are grouped in such a way as to support a certain decision or hypothesis. Instead of attempting to make a decision (detection or classification) in one step, either by a single sensor, or by a linear combination of a group of sensors, we propose to combine supporting evidence for a hypothesis in a hierarchical way by building a tree structure that combines at the lowest level clusters and in the next levels aggregates the outputs of several initial clusters in superclusters and so on. At each level in the tree decisions need to be made from different sources with different weights. This is conveniently modelled by the fuzzy λ-measure g_λ ($0 \leq g_\lambda \leq 1$). In particular we have in the absence of relevant information towards the classification/detection goal: $g(\emptyset) = 0$ and $g(A) \leq g(B)$ if $A \subseteq B$. This coincides with the intuitive feeling that if the evidence support is larger (*i.e.* if we observe the same scene with more sensors), that then the information content should also increase. In addition the following property holds for all $A, B \subset X$ with $A \cap B = \emptyset$:

$$\exists \lambda > -1 \;\; g(A \cup B) = g(A) + g(B) + \lambda g(A) g(B) \qquad (21)$$

This again supports the intuition that adding more independent data ($A \cap B = \emptyset$) co-operates towards an increase in confidence about the final decision. In addition both intuitive features about the fusion of two independent sensors are reproduced, *viz.* $\lambda \geq 0 \quad g(A \cup B) \geq g(A) + g(B)$, *i.e.* fusion is more than superposition and $-1 < \lambda \leq 0 \quad g(A \cup B) \geq MAX(g(A), g(B))$, implying that even if the synergy is negative (as reflected by the negative λ), it may still be fortuitous to apply sensor fusion. In the event that λ=0, *i.e.* the case where all sensors have the same importance and completely cover the universe of discourse, the degree of importance g_λ towards the final decision becomes additive and coincides with the definition of a probability measure.

Following ideas put forward in Ref. [23], sensor fusion may also be modelled using the concept of fuzzy integration. For a review on the role of fuzzy integrals in the framework of multiple criteria decision making see *e.g.* Ref. [24]. A fuzzy integral may be interpreted as an aggregation functional of subjective evidence, where the subjectivity is expressed in the fuzzy measure, and integration is defined over measurable sets [25]. In contrast to normal (Lebesque) integrals, fuzzy integrals are *non-linear* functionals. It is exactly this non-linearity and the possibility to include a fuzzy measure g_λ that is attractive in the context of fusion. Formally Sugeno's fuzzy integral is defined in the following way: Let X be a set of elements (e.g. sensors, features or classifiers) and let h(x): X→[0,1] denote the confidence value belonging to an element x∈X (e.g. the class membership of data determined by a specific sensor (classifier)), then the *fuzzy integral* of h(x) over a subset E of X with respect to the fuzzy measure g can be defined. The evaluation of the fuzzy integral may be interpreted as evaluating the degree of agreement between objective evidence h(x) and the expected observation outcome (the hypothesis). We will not discuss the properties of this fuzzy fusion operator here, but note that it is ideally suited to combine information from different sources *without* having to deal with the combinatorial explosion, as is the case in DS theory.

A totally different approach in finding a measure of synergy is followed in the present paper: In physics it is well-known that individual atoms, if joined together, produce new properties that are absent from the individual atoms. As an example we may consider the electric conductivity of a metal: one cannot define electric conductivity for isolated atoms. A metal has a bandstructure that originates from atomic levels of the individual atoms and symmetry breaking due to the Pauli principle. The idea that some properties only can exist for (large) ensembles of constituents, suggests that synergy may be modelled as such a property. Pursuing this idea further we noticed that emergent behaviour, *i.e.* the spontaneous occurrence of a characteristic of an ensemble may be modelled as an analogon to a phase transition in physics. The thermodynamics of phase transitions is very rich and also very non-linear.

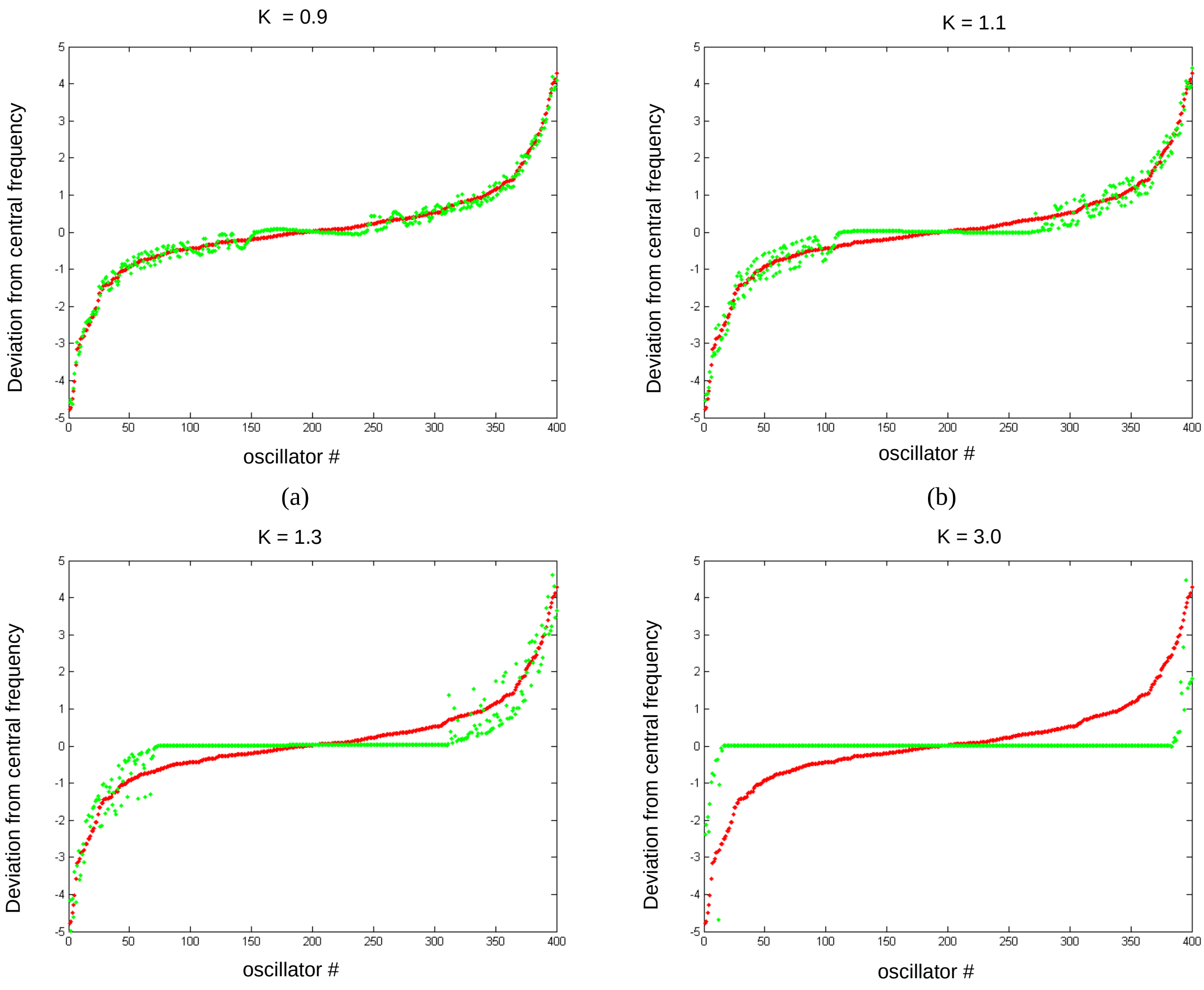
The discovery of Kuramoto showed that simple phase oscillators with all-to-all non-linear coupling display a second order phase transition if the interaction parameter is increased above its critical value, causing some of the oscillators to synchronize. This observation lead us to the idea that by mapping a scalar observation of a sensor as an eigenfrequency of an oscillator built into the sensor, the sensor is able to communicate (transmit) its data with its direct environment. If all sensors are equipped in this way and are also able to receive the results of neighbouring sensors, the proposed communication between the neighbouring nodes guarantees that via the synchronization phenomenon the sensors spontaneously co-operate and arrive at a common decision. there is no need to have a special node for sensor fusion: all sensor nodes are identical and it suffices to communicate with only one node to read out the result of the consensus. It is therefore clear that the Kuramoto synchronization phase transition can be viewed as a paradigm both for emergent collective behaviour, as well as a metaphor for modelling distributed sensor fusion, especially for global detection.

4. SIMULATIONS

In order to demonstrate the synergy induced by synchronization we carried out a number of preliminary simulations for $N = 400$ oscillators. The Kuramoto model, Eq.(10), with full connectivity between the nodes and equal weighting of all sensors $c_i=1$, was solved for different interactions K by Runge Kutta integration with a time step of 0.01 and at least 10^4 steps to guarantee stable solutions. We may, without loss of generality set the central frequency of the distribution $\omega^* = 0$. The eigenfrequencies of the oscillators, ω_i were randomly taken from a given distribution $g(\omega)$. We have taken a Lorentz distribution for $g(\omega) = \frac{\gamma}{\pi\left[\omega^2 + \gamma^2\right]}$, since for this particular case it is possible to obtain exact results in the thermodynamic limit, $N \longrightarrow \infty$ [see *e.g.* 16] . It follows from Eq. (15) that a minimum interaction strength, the critical strength K_C , exists above which a fraction of the oscillator ensemble will synchronize for $t \longrightarrow \infty$. In the case of the Lorentz distribution it is possible to find an analytical expression for the critical value: $K_C = 2\gamma$. The relation between r and the order parameter K above K_C is then given by: $r(K) = \sqrt{(1 - K_C / K)}$.

We have varied the interaction strength K between the oscillators from 0.7 to 3.0 in steps of 0.1 and displayed the results in Figs. 1 and 2. Note that in Figs 1a-d we have for reference also plotted the eigenfrequencies of the N oscillators. Because the eigenfrequencies were taken randomly with distribution function $g(\omega)$, the oscillators have been renumbered in ascending order according to their eigenfrequencies. From the results it is obvious that even below the critical interaction (Fig. 1a) already noticeable entrainment occurs. By increasing K further the fraction of synchronized oscillators also increases with the largest increase just above K_C as reflected by the order parameter r as defined in Eq.(6) (Fig. 2). From Fig.1 it is obvious that oscillators with eigenfrequencies close to the central frequency $\omega^* = 0$ synchronize more readily than oscillators with eigenfrequencies further away. This is compatible with the concept that the magnitude of the mean field initially is small, because the oscillators evolve incoherently in time. However above the critical strength (and in fact for finite N already just below the $N \longrightarrow \infty$ theoretical critical strength $K_C = 1$) some oscillators synchronize causing the mean field to build up.

Thus the synchronization phenomenon demonstrates positive feedback: the more oscillators become synchronized, the larger the mean field and the more oscillators become entrained. The synchronization process stops when the degree of coherence reaches a maximum value given by the interaction strength. If the interaction strength increases to very high values, the order parameter approaches 1, indicating that the whole population oscillates coherently. The oscillator phases behave in a similar way, although they need not converge to the same value. It should be noted that synchronization means that the time *derivatives* of the oscillator phases converge to the same value for $t \longrightarrow \infty$. The *phases* of the oscillators may vary depending on their initial conditions (Eq. (17)).



(c) (d)

Fig. 1a-d Distribution of frequencies of N=400 points of coupled phase oscillators with interaction strength K. The red dots indicate N random samples taken from the Lorentzian frequency distribution $g(\omega)=\gamma/\left[\pi(\omega^2+\gamma^2)\right]$ *with* $\gamma = 0.5$*, corresponding to the eigenfrequencies of the N oscillators. The oscillators have been ordered according to their eigenfrequencies. The central frequency coincides with* $\omega^* = 0$*. In the limit* $t \to \infty$ *the average deviations* $\dot{\vartheta}_k$ *from the central frequency* ω^* *are shown as green points. The interaction strength K is varied from 0.9 (a) through 1.1 (b) and 1.3 (c) to 3.0 (d), corresponding to just below the critical interaction* K_c*=1.0 to cases where synchronization occurs.*

In going from Fig. 1a to Fig. 1d we observe that the frequencies of the phase oscillators start to deviate from their “zero interaction” eigenfrequencies: some oscillators are already synchronized, others become “entrained” (pulled towards the central frequency), whereas still others move away from synchronization by increasing the deviation to the central frequency. If the interaction is increased, the fraction of synchronized oscillators increases as expected, but also the fraction of entrained oscillators increases at the cost of oscillators that move away from the central frequency. This is obvious in Fig. 1d where only a minute fraction of the oscillators has a frequency deviation from ω^* that is larger than the deviation of their eigenfrequency.

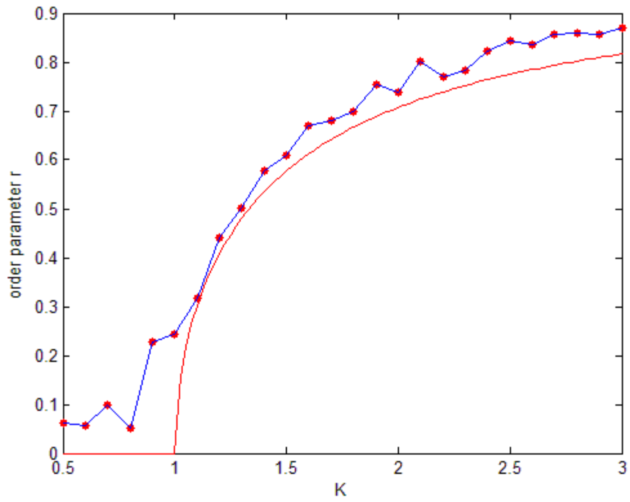


Fig.2 The order parameter r for $t \to \infty$ *as a function of the interaction strength K. The order parameter measures the fraction of the N oscillators that are synchronized. The red dots connected by the blue line are the result of calculations for N=400 oscillators and the same initial distribution of eigenfrequences as in Fig.1. It is seen that even below the critical interaction strength* K_c*=1.0 the order parameter r is not exactly zero, indicating that the onset of the phase transition for finite N is already noticeable below the critical interaction strength. Also inserted is the theoretical phase boundary in the thermodynamic limit* $N \to \infty$*, given by r=0 for* $K < K_c$ *and* $r = \sqrt{(1-K^{-1})}$ *for* $K \geq K_c$.

5. CONCLUSION

The preliminary simulation results show that a system of interacting phase oscillators with non-linear all-to-all interaction (*i.e.* an ensemble of agents defined on a fully connected graph) is capable of producing a collective property, *viz.* synchronization. Synchronization can therefore be used as a paradigm for emergent behaviour, because it originates spontaneously at certain critical interaction strength. The transition from an asynchronous ensemble to a coherent state where a large fraction of oscillators is synchronized can be described as a phase transition in thermodynamics. In terms of a phase transition, we may describe this phenomenon with a so-called order parameter (the quantity r in our notation). Sensor fusion can benefit from this paradigm in a number of ways. In the first place synchronization constitutes a form of physical layer communication between the sensors and in the second place it is possible to model consensus (arriving at a common decision, *e.g.* in detection) using the phenomenon of phase synchronization. The effectiveness of sensor fusion, *i.e.* arriving at consensus on detection can therefore be characterized with this order parameter. As soon as the non-linear interaction strength is increased above a certain value, or alternatively, if the surface density of sensors (*e.g.* in smart dust) is increased above a certain threshold, suddenly co-operation emerges that helps establish a *"communis opinio"* among the interacting oscillators. This happens suddenly, not gradually, and therefore cannot be described by a superposition of individual sensor contributions. It is clear that this phenomenon is extremely advantageous because it is simultaneous and fast, without the need for intricate protocols, no special fusion centre is necessary and routing to a fusing centre is not needed. Moreover it appears that the overall robustness of the network is thus increased because the fusion process is carried out by the network *as a whole* and not by highly specialized centres. Finally we may conclude that non-linear interactions open avenues towards a quantitative basis of intuitive notions such as the added value of synergy.

6. REFERENCES

[1] P. Gupta and P.R. Kumar, J. Sel.Areas Commun. (Special issue on self-organizing distributed collaborative sensor networks), vol 23, 4 pp. 755-764 (2005).

[2] A. Giridhar and P.R. Kumar, IEEE Trans. Inf. Theory vol 46, 2 pp. 388-404 (2000).

[3] P. Horowitz and W. Hill, The art of electronics, 2nd edition, Chapter 5, Cambridge University Press, Cambridge (1989).

[4] L.O. Chua, C.A. Desoer, and E.S. Kuh, Linear and non-linear circuits, McGraw-Hill, New York, pp. 431-439 (1987).

[5] A. Liénard, Revue générale de l'électricité **23**, pp. 901–912 and 946–954 (1928).

[6] S. Lefschetz, Differential Equations : Geometric Theory, Dover Publications, New York, Chapter 11 (1977).

[7] N. Minorski, Non-linear oscillations, R.E. Krieger, Malabar, Florida, Chapter 20 (1962).

[8] L. Perko, Differential Equations and Dynamical Systems, Springer, New York, pp. 250-261 (1996).

[9] F. Verhulst, Nonlinear differential Equations and Dynamical Systems, Springer, New York (1989).

[10] S. Barbarossa and F. Celano, Proc. IEEE Conf. on Signal Processing Advances in Wireless Communications, pp. 475-479 (2005).

[11] Y. Kuramoto, Int. Symposium on Mathematical Problems in Theoretical Physics, vol 39 pp. 420, Springer, New York (1975).

[12] J.A. Acebrón, L.L. Bonilla, C.J. Pérez Vicente, F. Ritort, and R. Spigler, Reviews of Modern physics, vol 77, pp. 137-185 (2005).

[13] R.E. Best, Phase-Locked Loops, 6th edition, McGraw-Hill, New York (2007).

[14] N.N. Bogoliubov and Y.A. Mitropolsky, Asymptotic Methods in the theory of non-linear oscillations, Hindustan Publishing Corp., New Delhi (1961).

[15] N.I. Margaris, Theory of the non-linear analog phase locked loop, Springer (2004).

[16] A. Pikovsky, M. Rosenblum, and J. Kurths, Synchronization, a universal concept in non-linear sciences, Cambridge University Press, Cambridge (2001).

[17] G. Resconi and A.J. van der Wal, Information Sciences 142, pp. 249-273 (2002).

[18] T. Bayes, Philos. Trans. Royal Soc. London, vol 53, pp. 370-418 (1763).

[19] A.P. Dempster, Ann. Math. Statistics 38, pp.325-339 (1967).

[20] G. Shafer, A mathematical theory of evidence, Princeton University Press, Princeton (1976).

[21] S.C.A. Thomopoulos, J. Robotic Syst. Vol 7(3), pp. 337-372 (1990).

[22] M. Sugeno, Fuzzy automata and decision processes, North Holland, Amsterdam, pp. 89-102 (1977).

[23] J.M. Keller, h. Qiu, and H. Tahani, Proc. NAFIPS June 1986, pp. 324-338 (1986).

[24] M. Grabisch, IEEE Trans. Fuzzy Syst. 3(1), pp. 96-109 (1995).

[25] W.F. Pfeffer, Integrals and measures, Marcel Dekker, New York (1977).