The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

2052-50

Summer College on Plasma Physics

10 - 28 August 2009

Information theory and plasma turbulence

Richard Dendy

Euratom/UKAEA Fusion Association
Culham Science Centre
UK

Information theory and plasma turbulence

Richard Dendy

Euratom/UKAEA Fusion Association
Culham Science Centre
Abingdon, Oxfordshire OX14 3DB, U.K.

Centre for Fusion, Space and Astrophysics
Department of Physics, Warwick University
Coventry CV4 7AL, U.K.

Work supported in part by UK Engineering and Physical Sciences Research Council

Tomaso Aste’s (non-plasma-based) criteria for complex systems

Complex systems

- exhibit **emergence:** some properties present at system level are not present at lower level — e.g. a cell is alive but is made of inanimate elements
- are **open:** energy and information are constantly being imported and exported across system boundaries
- have a **history:** the history cannot be ignored, even a small change in circumstances can lead to large deviations in the future
- can **adapt:** in response to external or internal changes, the system can reorganize itself without breaking — self organising
- are not completely **predictable:** when a system is adaptive, unexpected behaviours can emerge — prediction becomes expectation
- are **multi-scale** and **hierarchical:** system size and structure scale are over several orders of magnitude and distinct properties and functions are associated with different scales; dynamics can propagate through scales — avalanches, cascade effects
- are **disordered:** there is no compact and concise way to encode the whole information contained in the system
- have **multiple** (meta) (stable) **states:** small perturbations lead to recovery, larger ones can lead to radical changes of properties; dynamics do not average simply

http://wwwrsphysse.anu.edu.au/~ccs106/SUMMERSCHOOLS/SS22/Proceedings/Themes.shtml

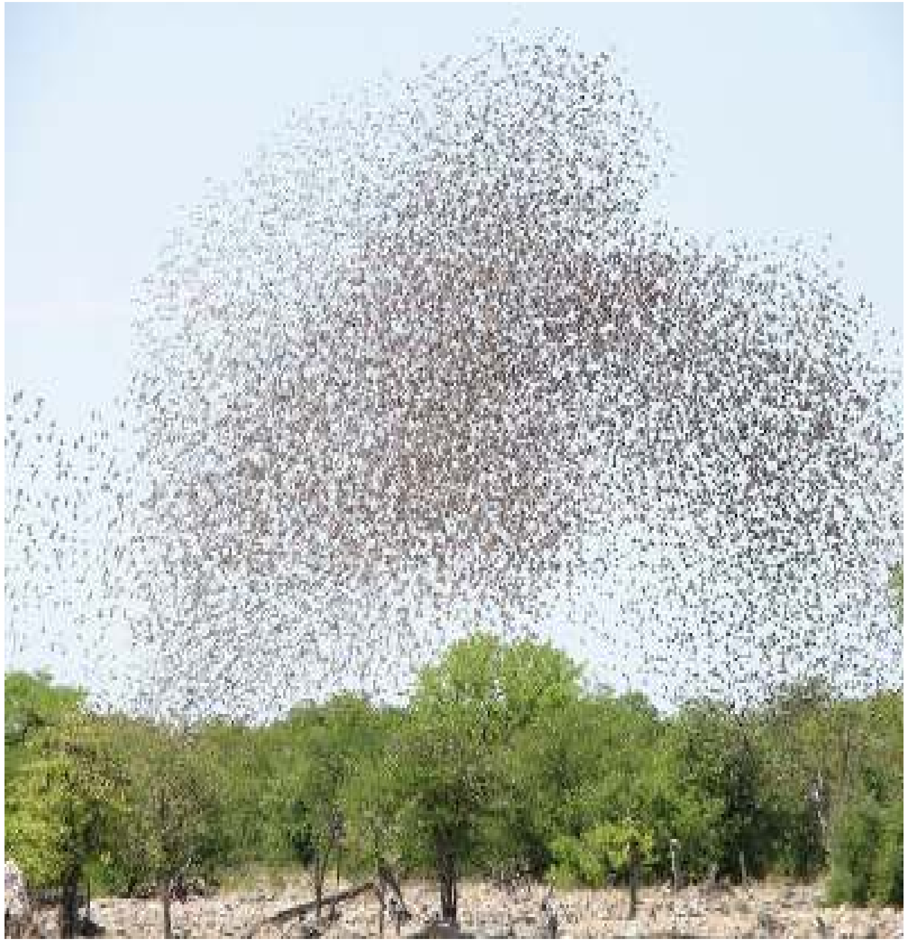
CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe

A word from our feathered friends

Emergent self organisation, or what? Type “starlings Otmoor” into Google

Quelea



Starling

Aste's complex systems criteria as applied to plasmas (1)

- Plasmas in fusion and space exhibit *emergence*, meaning that *some properties present at system level are not present at lower level*.

- Overall energy confinement is a property that emerges only at *system level* from the interplay of *coupled* physical processes operating across a *hierarchy* of lower levels, reaching down to single particle dynamics.

- Each level of description (single fluid; two fluids – electrons and ions; kinetic ions and fluid electrons; gyrokinetic; and so on) within this hierarchy is determined by the characteristic lengthscale and timescale of whichever physical process dominates at that level. Plasmas are thus *multiscale*.

- The different levels of description and associated observed phenomenology *extend over several orders of magnitude, and distinct properties and functions are associated with different scales*.

- Plasmas self organise persistent coherent macroscopic structures that only arise on lengthscales at, or just below, *system level*. Examples include magnetic islands and zonal flows, which *are not present at lower level*.

Aste's complex systems criteria as applied to plasmas (2)

- Plasmas are invariably *open*, in the sense that *energy and information are constantly being imported and exported across system boundaries*.
- The quest for fusion power from magnetically confined plasmas involves injecting energy at the 10 MW level into a gram of material occupying a volume of tens of cubic metres. That such plasmas sustain, over seconds, the steepest steady-state temperature gradients known, while subject to energy fluxes of several MWm^{-2}, shows their ability to *adapt*: *in response to external or internal changes, the system can reorganize itself without breaking*.
- Plasmas are *not completely predictable*: *unexpected behaviours can emerge – prediction becomes expectation*. Performance in future experiments is extrapolated using empirical dimensionless scaling laws in the absence of first principles predictions of global phenomenology. Key behaviours, such as enhanced confinement operating regimes and ELMs, were not predicted.
- Transitions between confinement regimes typically have a *history*: *even a small change in circumstances can lead to large deviations in the future*, and reflect the existence of *multiple metastable states*. These transitions can occur spontaneously as plasma conditions evolve in time, or can be induced by careful sequencing of external drivers, notably auxiliary heating and fuelling. History is crucial and there is an element of irreversibility.

And so to information…

- The typical very low density and high temperature of plasmas in magnetically confined fusion experiments implies a high degree of *disorder* at the lowest level of description, namely the self-consistent dynamics of charged particles and electromagnetic field.
- For this reason, *there is no compact and concise way to encode the whole information contained in the system:*
  - particle-in-cell codes which implement this lowest-level description are best adapted to phenomena occurring on the fastest timescales and shortest lengthscales
  - higher level descriptions are reduced models; to construct these, information has deliberately been dropped.
- So, for plasmas and for complex systems in general, what is the *information contained in the system*?

This makes information theory for plasmas topical and exciting

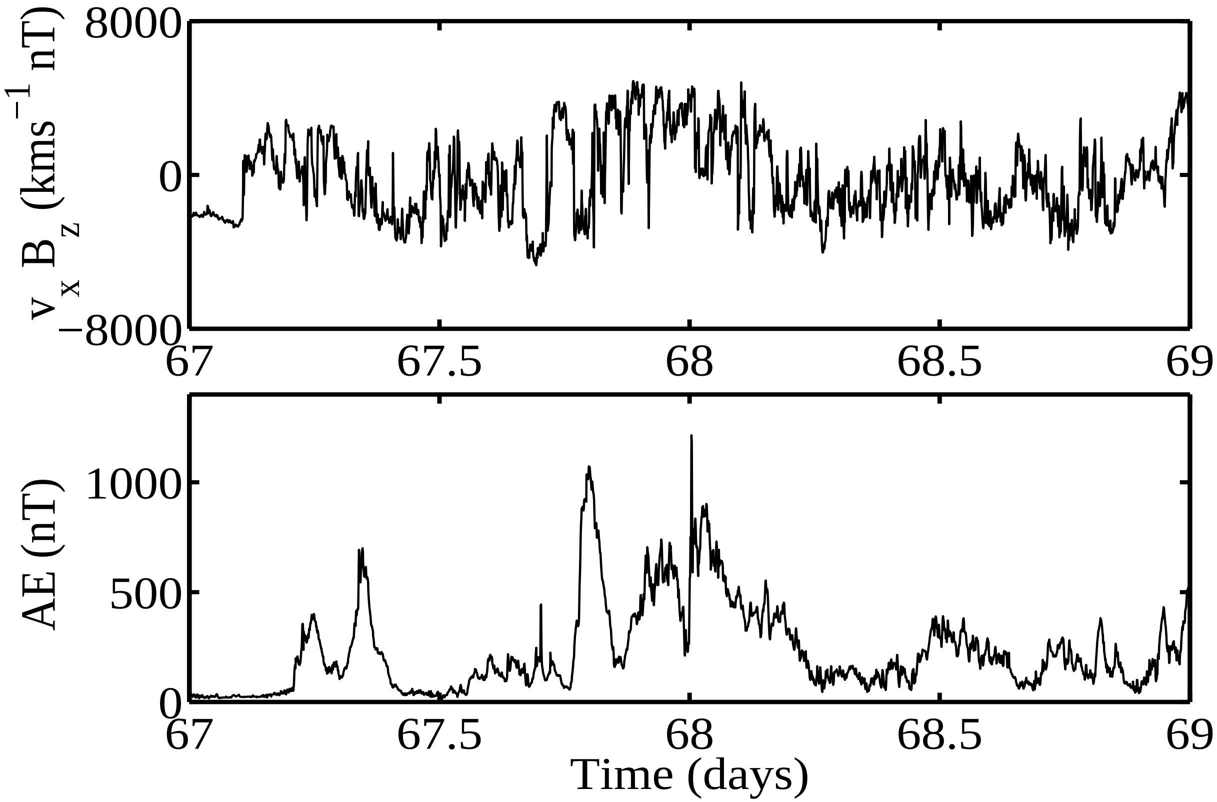
It appears that information theory may in future provide central unifying principles for complex systems science and plasma physics alike

- Irrespective of their physical and mathematical embodiment, all complex systems have in common the creation, transmission, sharing and destruction of information.
- It is the ebb and flow, birth and death of information – a physical quantity – that underlies and enables the physical phenomenology.
- Quantifying the state and distribution of information within a complex system is thus crucial both to understanding its working, and to rigorously characterising its behaviour.
- We outline some pioneering studies of mutual information in the solar wind plasma upstream of Earth, using techniques tested on standard complex systems models for the collective dynamics (i.e., flocking) of birds.
- A fundamental question of great interest arises. Namely, the relation between:
  - information evaluated at the bulk level of description – for example, mutual information used as a measure of nonlinear correlation between spatiotemporally separated but causally linked fluid flows; and
  - information evaluated when finer structure – for example, in measurements of magnetic field fluctuations – is resolved in the same system.

Quantifying physical linkage of two spatiotemporally separated, highly nonlinear, plasma signals: upstream solar wind and ionospheric

Solar wind from WIND satellite at sunward libration point

Terrestrial magnetometer data at high geomagnetic latitude



Solar wind drives magnetotail reconnection: energy release drives ionospheric currents affecting terrestrial magnetic field

What is information?

Information resides in the number of yes/no (≡ binary 0/1) questions (≡ *bits*) to which we have the answer. E.g. for n = 3 questions there are:

$2^3 = 8$ possible combinations of yes/no answers, expressible as

8 three-digit (≡ three-bit) binary *symbols* 101, 110,…, etc.,

So in general *n bits* → an *alphabet* containing $M = 2^n$ *symbols*

Suppose we *sample* (ask n questions) the system on N occasions. The amount of *information* thereby obtained, H, is the number of questions to which we have answers:

$$H = N \times n = N \log_2 M$$

If all symbols occur with equal statistical probability $P = 1/M$, then

$$H = - N \log_2 P$$

Any digitally sampled measured signal is a time-ordered string of N n-bit symbols $X_1, X_2, .., X_i, \ldots, X_N$ drawn from an alphabet having $M = 2^n$ symbols.

Different symbols X_i recur N_i times, implying different empirical probabilities

$$P_i = N_i/N \neq 1/M$$

CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe

Information and signal measurement

Intuitively, the occurrence in the signal of a statistically rare symbol (small P_i, e.g. letter “x”) provides more information H than the occurrence of a frequent one (large P_i, e.g. letter “e”)

For the equal probability case, we also know $H = - N \log_2 P$ for N symbols, implying information per symbol = $H/N = - \log_2 P$

It appears logical to define the information gained from a single occurrence of X_i as $- \log_2 P_i$

In the signal of length N symbols, X_i occurs N_i times. So the total information provided by the occurrences of X_i is $H_i = - N_i \log_2 P_i$

The total information in the signal is then

$$H = \sum_i H_i = - \sum_i N_i \log_2 P_i = - \sum_i N P_i \log_2 P_i = - N \sum_i P_i \log_2 P_i$$

Hence the average information per symbol in a real signal is

$$h = H/N = - \sum_i P_i \log_2 P_i$$

This is the *Shannon Entropy* of the signal: “entropy” because of deep analogies with statistical mechanical entropy and, beyond, to thermodynamic entropy

Relevance of information theory to complex systems science

Complex systems typically yield highly nonlinear measurements – intermittent, bursty

Hence it may be suboptimal to try to identify correlation and causality via Fourier-derived techniques that rest upon the superposition of linear modes

Information-based analysis is intrinsically nonlinear, being based on sets of probabilities of arbitrary relative magnitude

The strategy is:

- Split each measured signal into a time-ordered string of symbols X_i
- Bin the data symbols to establish their probabilities P_i
- Calculate how information (meaning $-\sum_i P_i \log_2 P_i$ type quantities) is shared, flows, and decays, both
    - within a given signal
    - between two contemporaneous but separate signals

These techniques are widely used but remain "novel" across a broad range of physics, including complex systems science and plasma physics

Defining linear cross covariance and mutual information

Both provide measures of correlation between two signals *A* and *B*.

Linear cross covariance

$$C(A,B) = \frac{E[(A-\overline{A})(B-\overline{B})]}{\sqrt{E[(A-\overline{A})^2]E[(B-\overline{B})^2]}}$$

where *E* [...] denotes the mathematical expectation value, and $\bar{A} = E[A]$.

Nonlinear mutual information

$$I(A,B) = \sum_{i,j}^{m} P(a_i,b_j)\log_2\left(\frac{P(a_i,b_j)}{P(a_i)P(b_j)}\right)$$

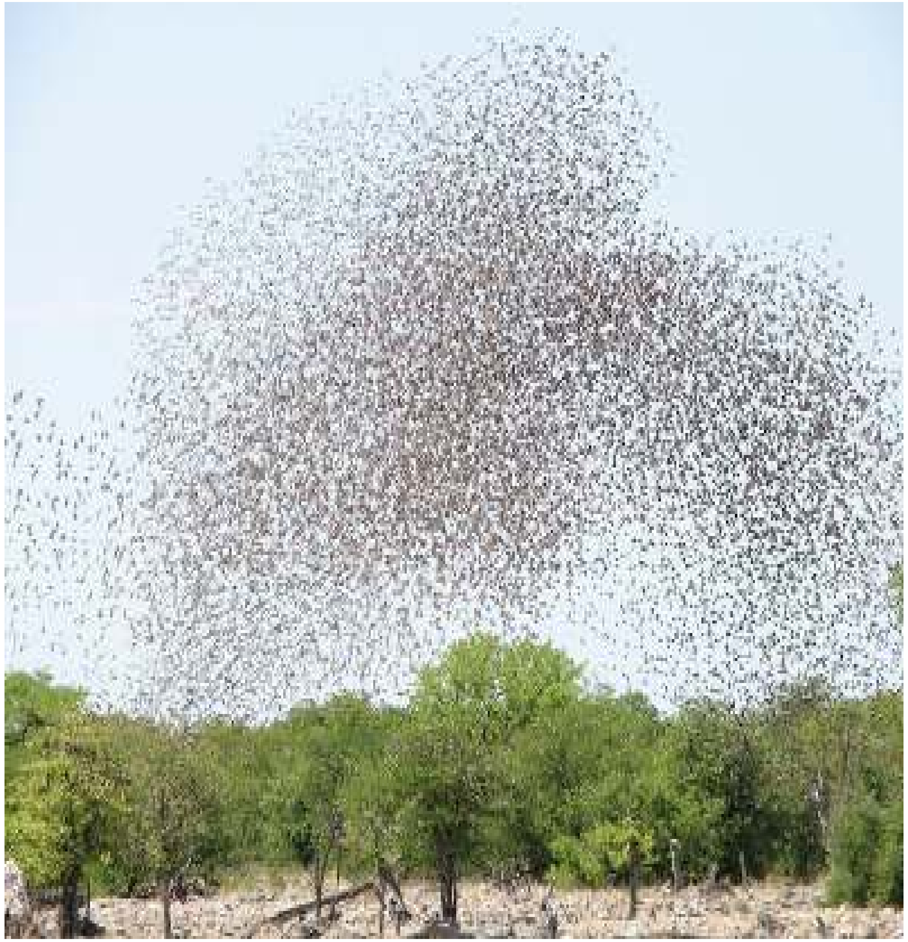
where signals *A* and *B* have been partitioned into exhaustive discrete alphabets $\{a_i\}$, $\{b_j\}$, with

- each symbol having empirical probability $P(a_i)$, $P(b_i)$,
- $P(a_i,b_j)$ is the joint probability of a_i and b_j

A word from our feathered friends

Emergent self organisation, or what?

Quelea



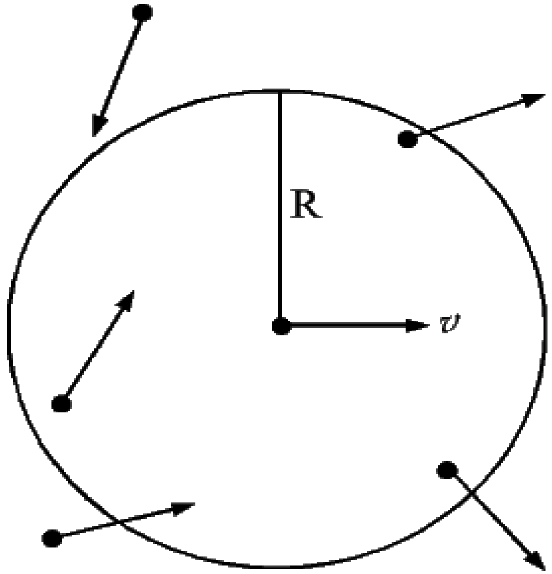
Starling

Quantifying clumpiness and flocking in complex systems, including plasmas

The Vicsek model* for flocking birds, fish,...

-Each flying bird (swimming fish...) takes account of the velocity orientation of its near neighbours, and does its best (subject to noise) to align with them

-Speed is constant, velocity orientation and position change



For each bird, at each successive time step:

-update position using current velocity

-identify the other birds within radius R, take their average velocity orientation, and add noise

$$x_{n+1} = x_n + \vec{v}\,\delta t$$

$$\theta_{n+1} = \langle \theta_n \rangle_R + \delta\theta_n$$

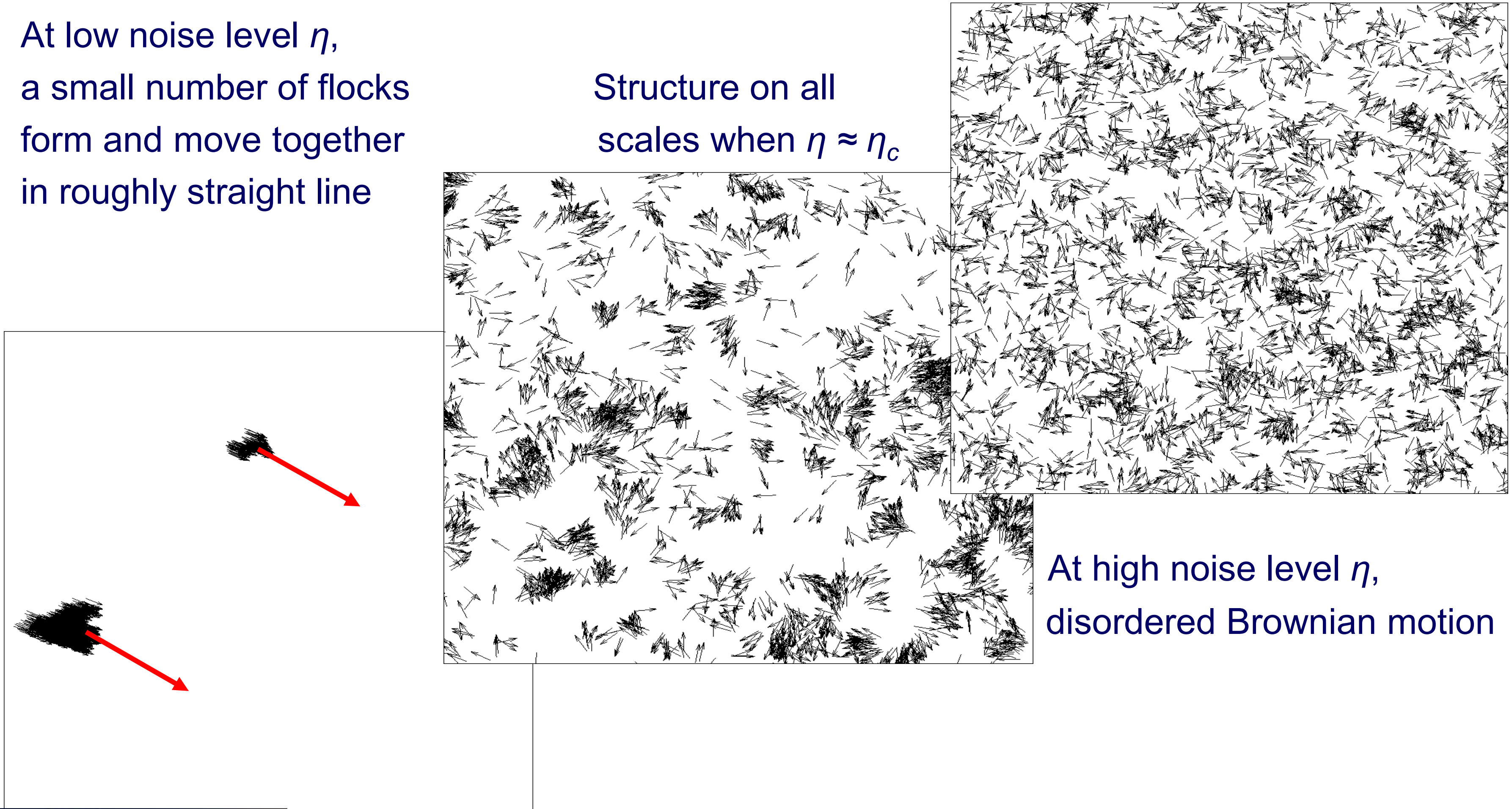
Noise range is $-\eta < \delta\theta < \eta$

Vicsek, Czirok, Ben-Jacob, Cohen & Shochet, Phys. Rev. Lett. ***75****, 1226 (1995)*

Critical phase transition at noise $\eta = \eta_c$ in the Vicsek model

At low noise level η,
a small number of flocks
form and move together
in roughly straight line

Structure on all
scales when $\eta \approx \eta_c$



At high noise level η,
disordered Brownian motion

CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe

Quantifying the phase change in the Vicsek model

Classical physics measures are
- "Order parameter" φ, in this case mean velocity

$$\phi = \frac{1}{N v_0} \left| \sum_{i=1}^{N} v_i \right|$$

- "Susceptibility" χ, in this case velocity dispersion

$$\chi = \sigma^2(\phi) = \frac{1}{N}\left(\langle \phi^2 \rangle - \langle \phi \rangle^2 \right)$$

Information theory measure is derived from the probability distribution of the birds' positions and velocities $\{x_n, \theta_n\} \equiv A \equiv \{a_1, a_2, a_3, \ldots\}$: the "signal" comprising the "alphabet" (i.e., pre-assigned set of strings) a_i, each of which is found to occur with measured probability $p(a_i)$

From these probabilities we can construct the Shannon entropy

$$H(A) = -\sum_{i=1}^{n} P(a_i) \log_2 (P(a_i))$$

Given two such signals, we can measure their information theoretic correlation in terms of their normalised mutual information

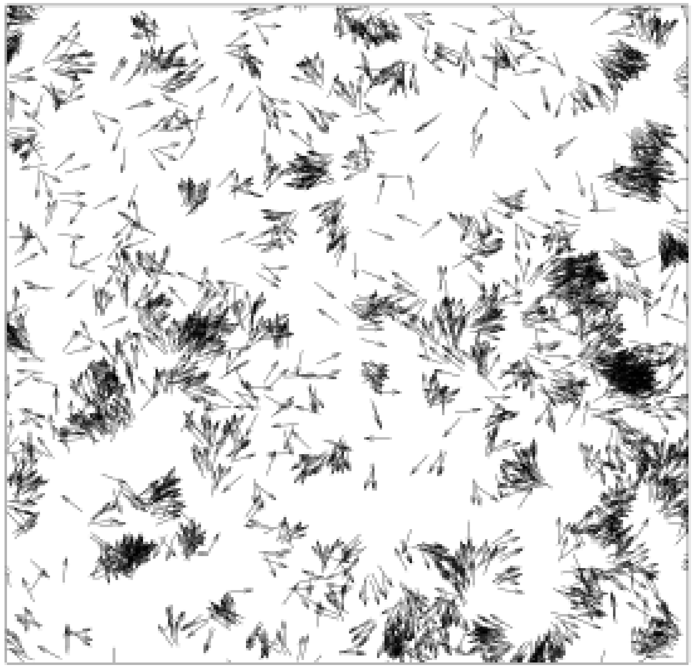
$$NMI(A, B) = \frac{H(A) + H(B)}{H(A, B)} - 1$$

CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe

Strategy for calculating mutual information for Vicsek system

1. Take a snapshot



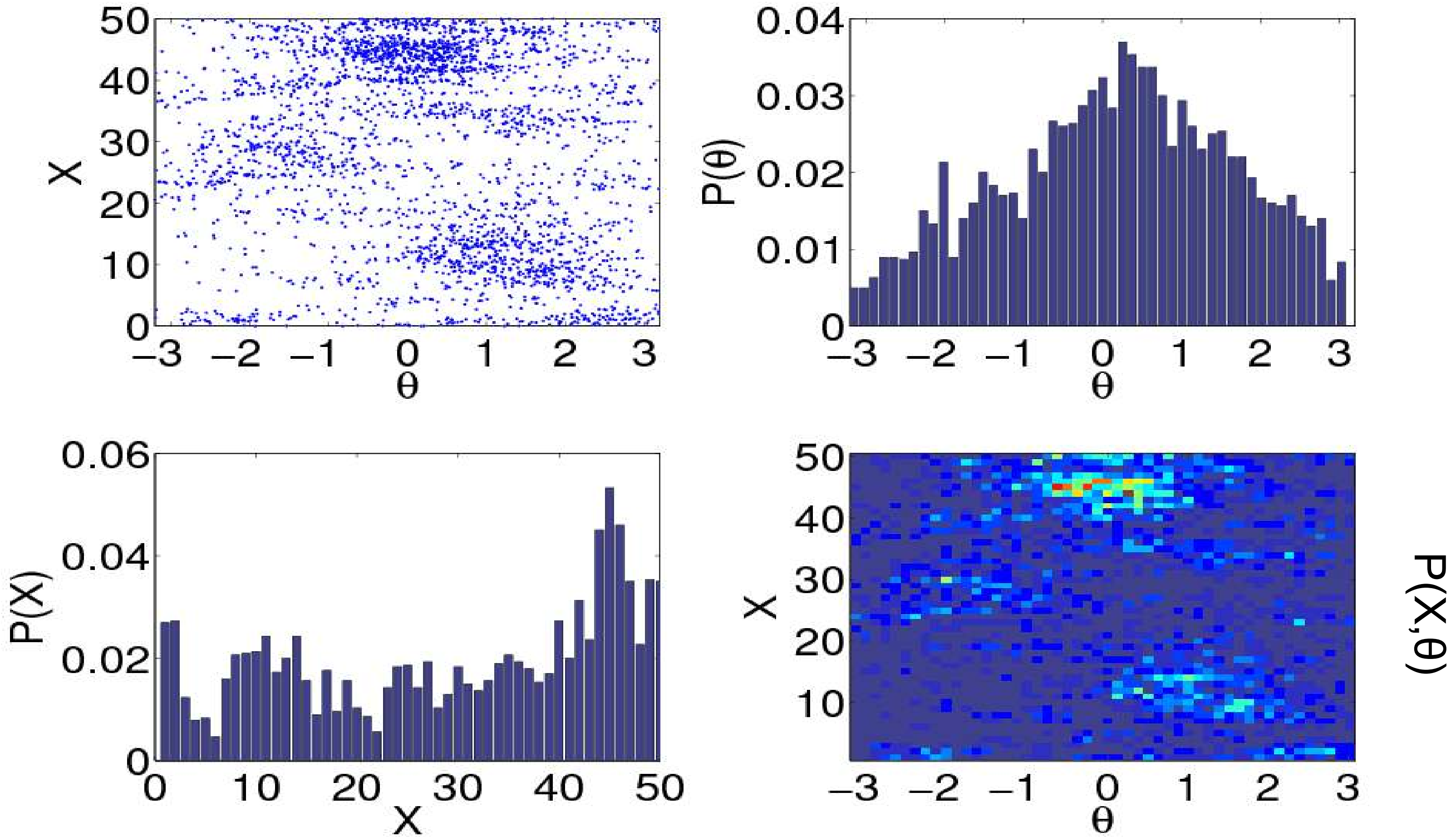
2. Discretise data – positions *x* and velocity orientations *θ* work best

3. Calculate entropies

Measuring mutual information in the Vicsek system

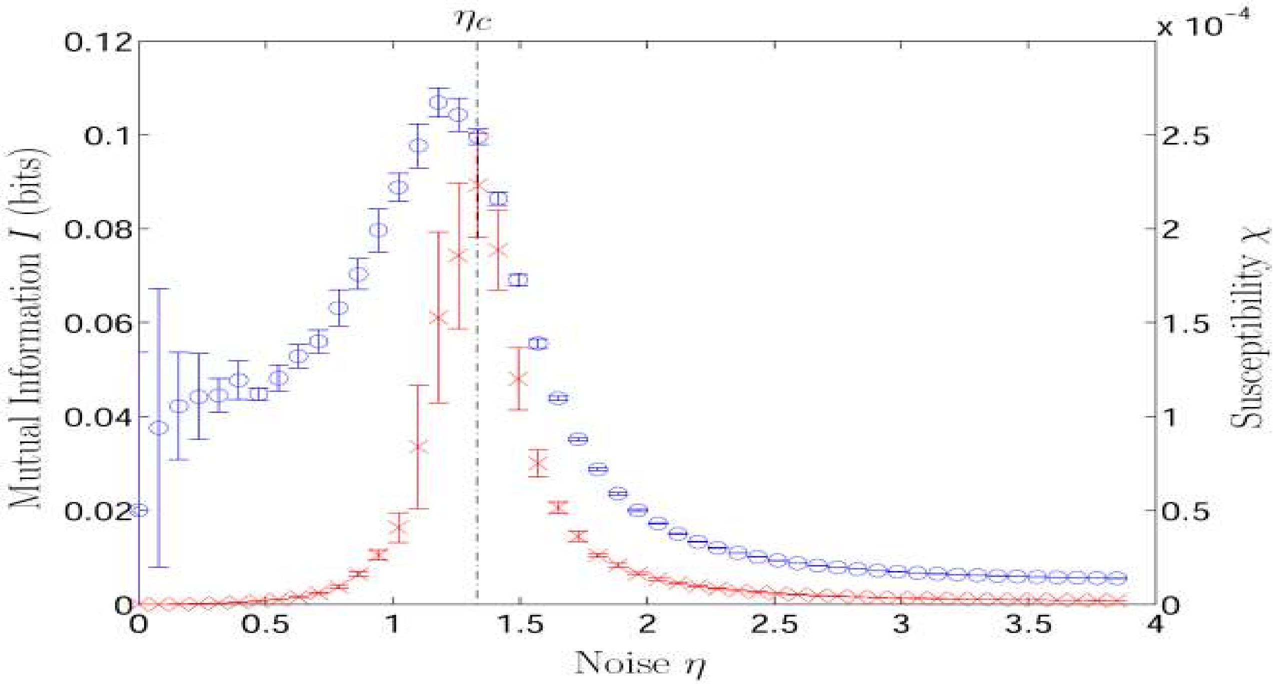
Move from actual distributions to P(x), P(θ), and P(x, θ)

Raw data



Classical and information theory results for Vicsek

Plot measured mutual information (blue) and susceptibility (red) versus noise η



Error bars near the peak identifying the phase transition are

- at their smallest for mutual information

-at their largest for susceptibility

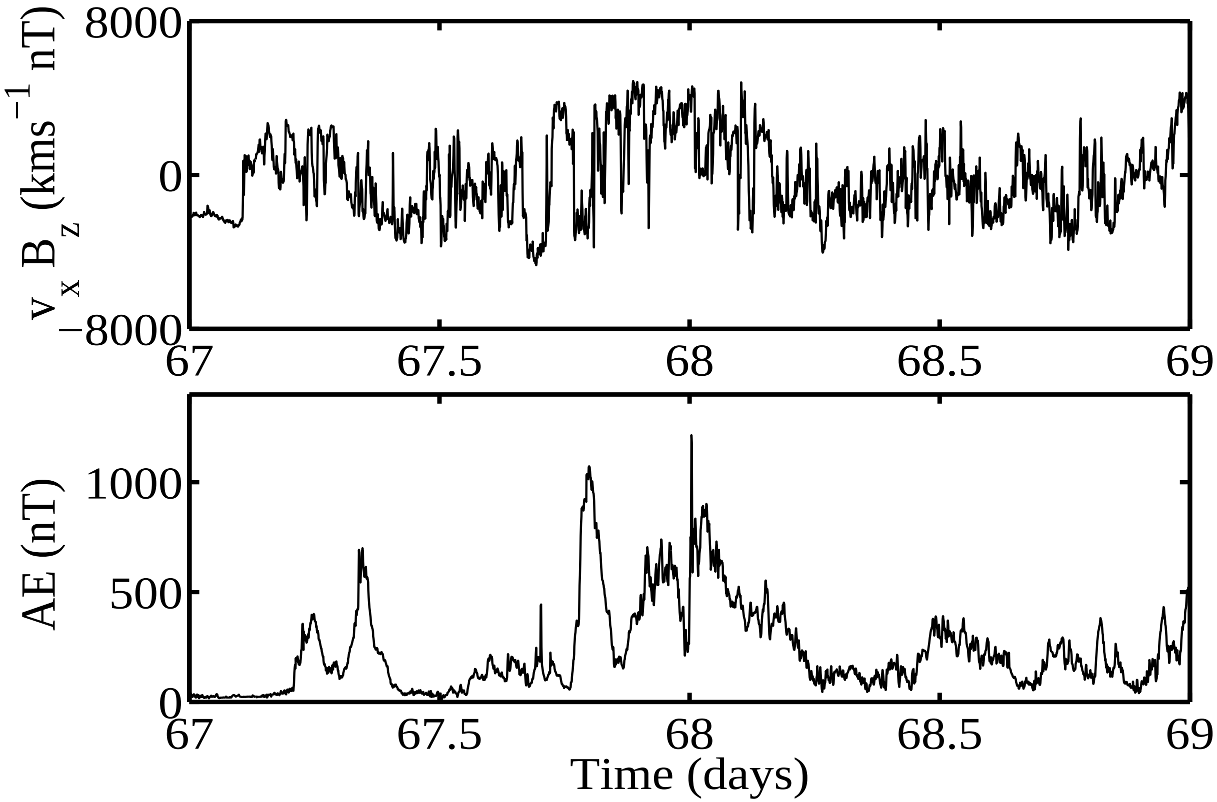
In this respect the intrinsically nonlinear measure is "better"

Wicks, Chapman & Dendy, Phys. Rev. E **75, 051125 (2007)*

Quantifying physical linkage of two spatiotemporally separated, highly nonlinear, plasma signals: upstream solar wind and ionospheric

Solar wind from WIND satellite at sunward libration point

Terrestrial magnetometer data at high geomagnetic latitude



Solar wind drives magnetotail reconnection: energy release drives ionospheric currents affecting terrestrial magnetic field

CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe

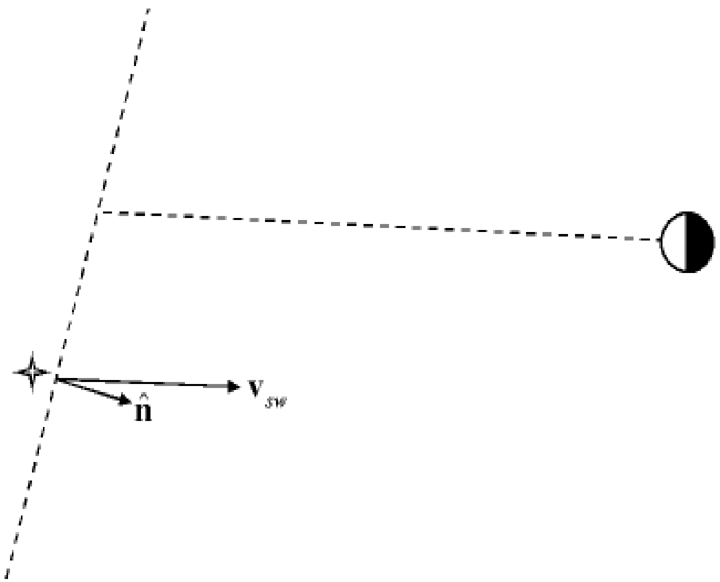
How much information do the solar wind and magnetometer data share in common?

March, Chapman and Dendy, *Geophysical Research Letters* **42**, L04101 (2005)

- Distinguish between hypotheses concerning solar wind propagation
  - Project ST data series in time according to different hypotheses for $\mathbf{v}_{sw}$ and $\mathbf{n}$:

    $\Delta t = (\mathbf{P}_w - \mathbf{P}_E) \cdot \mathbf{n}/\mathbf{v}.\mathbf{n}$

- Time lag introduced by magnetospheric plasma processes
  - Additional $\Delta t'$ to accommodate this



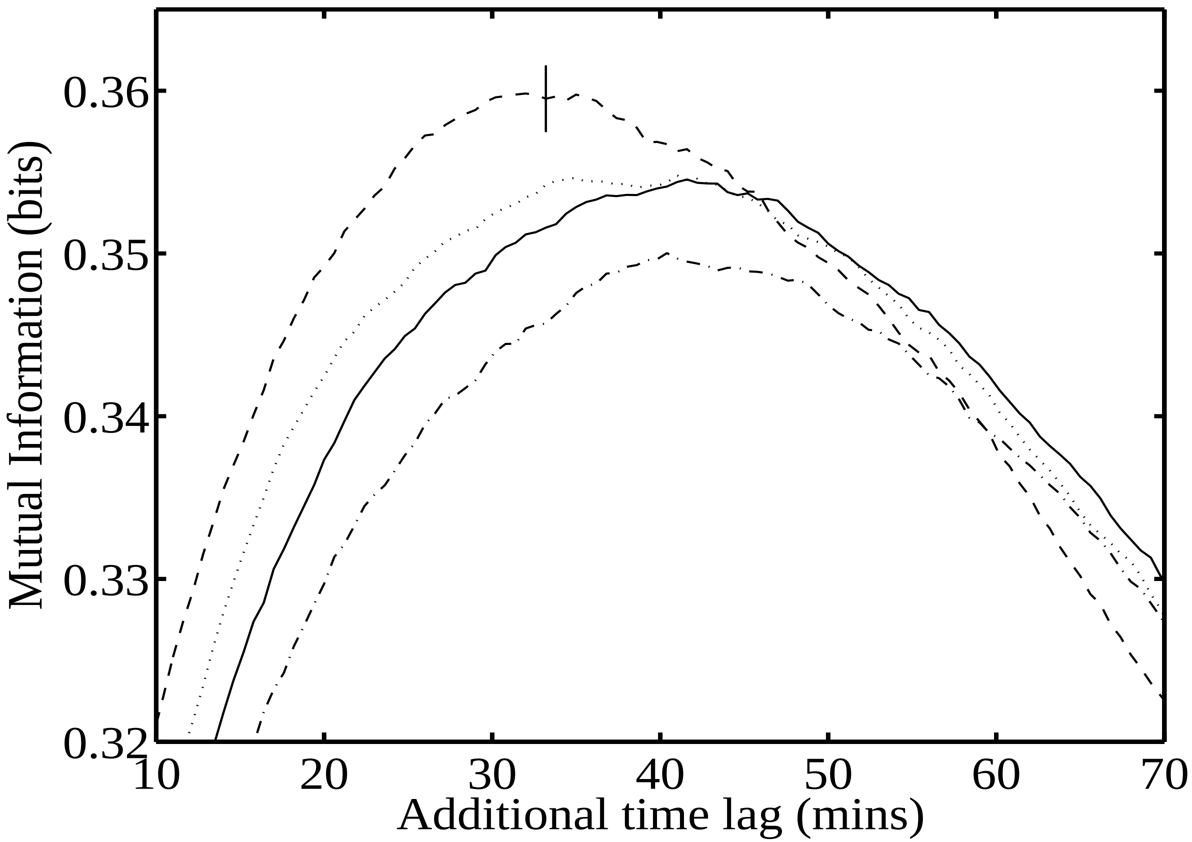
- Compute mutual information between SW(t) and AE $(t + \Delta t + \Delta t')$, and maximise

Mutual information between AE and SW

For four different SW propagation hypotheses, as a function of additional time lag $\Delta t'$

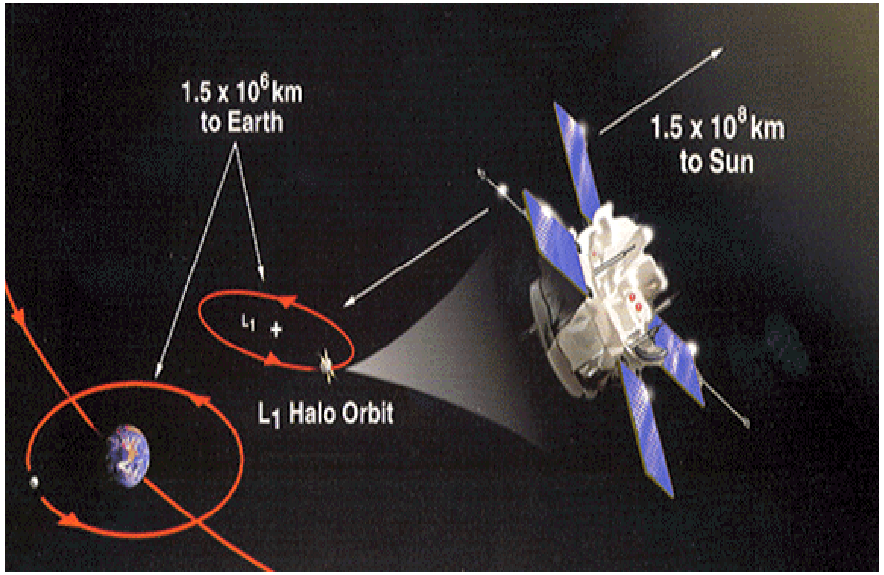
Physics output:

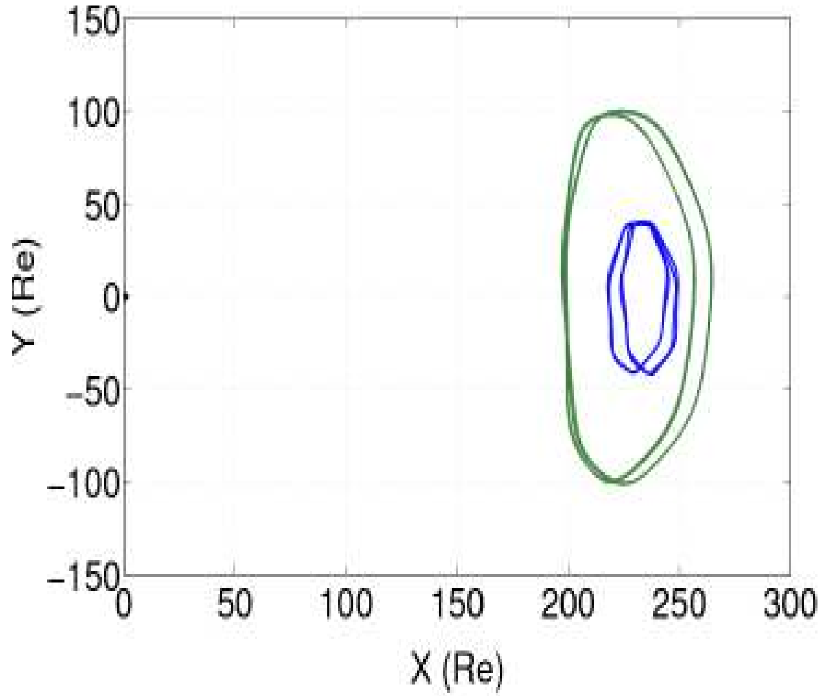
- Best hypothesis for propagation
- Shared information quantified
- Best time lag in magnetosphere



Simultaneous two-spacecraft measurements in the solar wind

During 2005-2006 the WIND and ACE spacecraft were at the Sun-Earth libration point L_1 in the distant upstream solar wind – a near-ideal (remote boundaries, broad range of scales) turbulent plasma





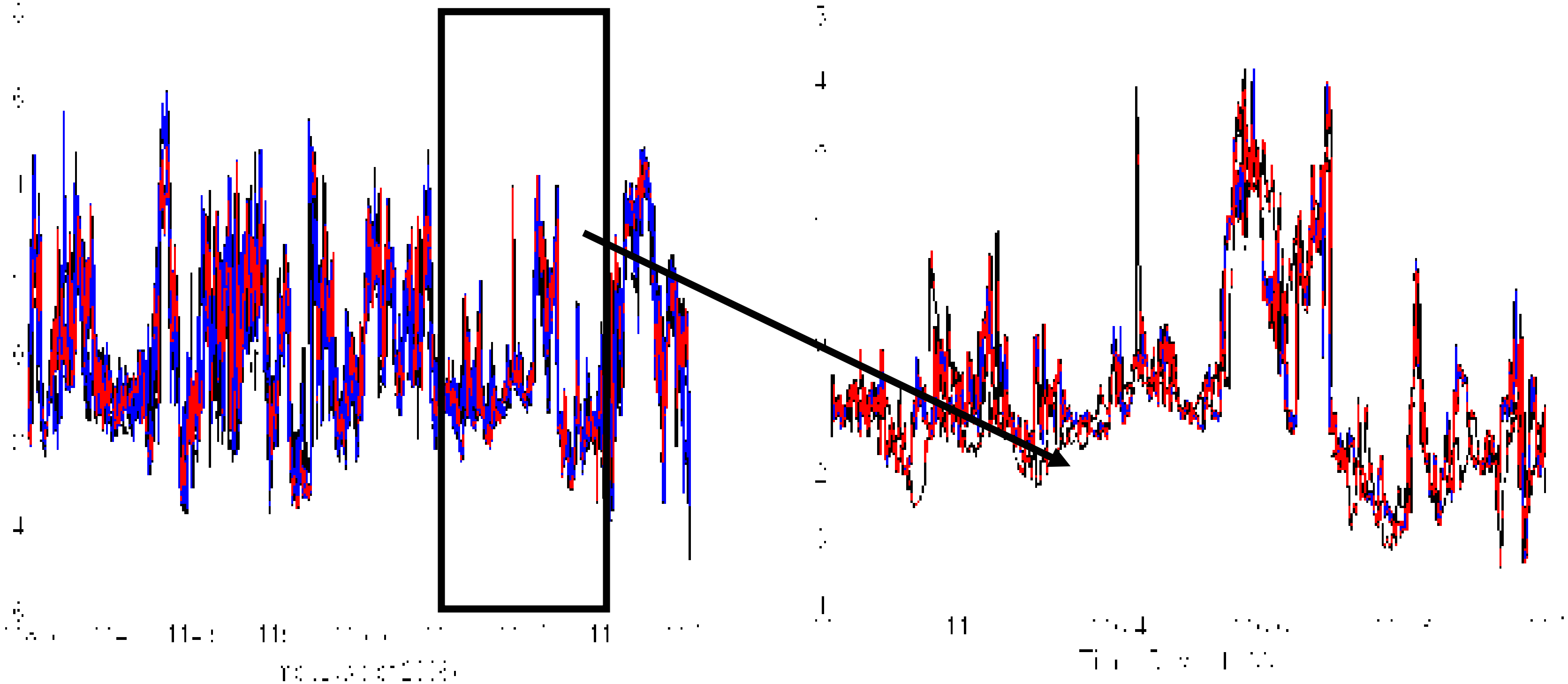
This enables measurements of spatial correlations in the measured, highly nonlinear, plasma and magnetic field properties of the solar wind, over a separation range 30 to 100 Earth radii (Re)

CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe

Simultaneous solar wind measurements from WIND and ACE

Typical observations of B_x [nT] from WIND (blue) and ACE (red)

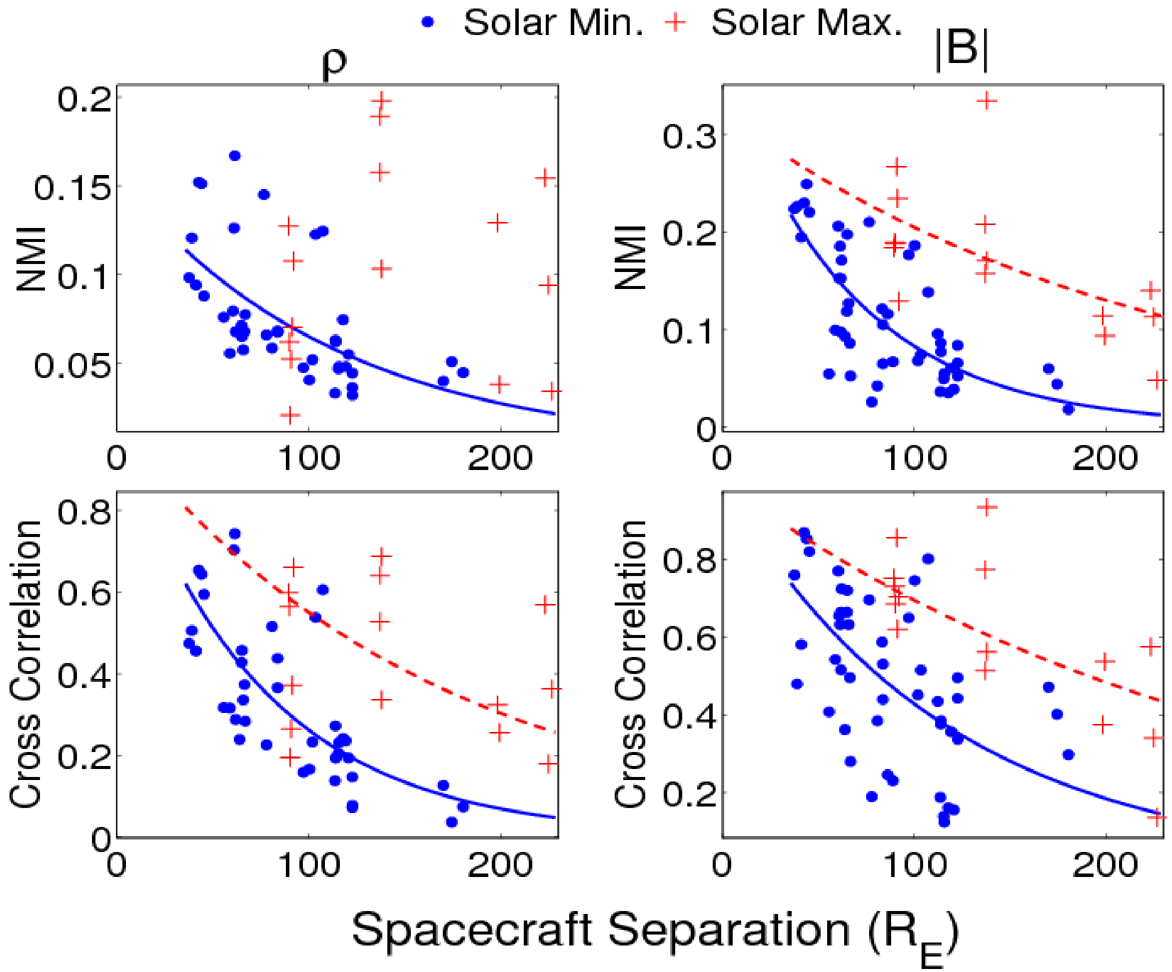


Four day trace | Embedded one-day trace

Strongly nonlinear signals exhibiting correlation across a range of timescales

Mutual information measures the spatial decay of correlation

Density Magnetic field strength



Mutual Information as described

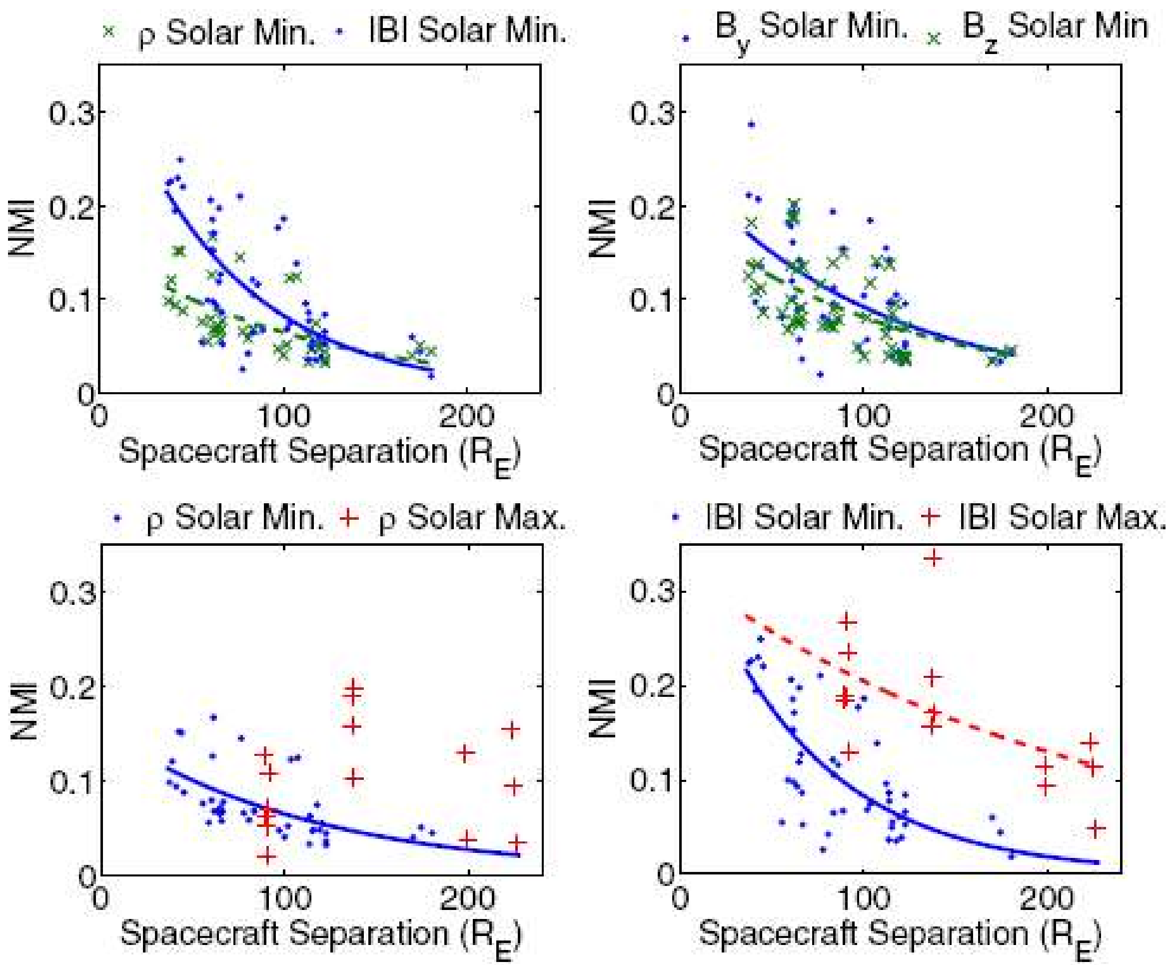
Cross correlation

$$C(A,B) = \frac{E[(A-\overline{A})(B-\overline{B})]}{\sqrt{E[(A-\overline{A})^2]E[(B-\overline{B})^2]}}$$

Wicks *et al,* Astrophys J **690** 734 (2009)

Fitted curves of exponentially decaying shared information help distinguish the physical nature of of the nonlinear structures

Spatial decay of correlation depends on solar activity in different ways



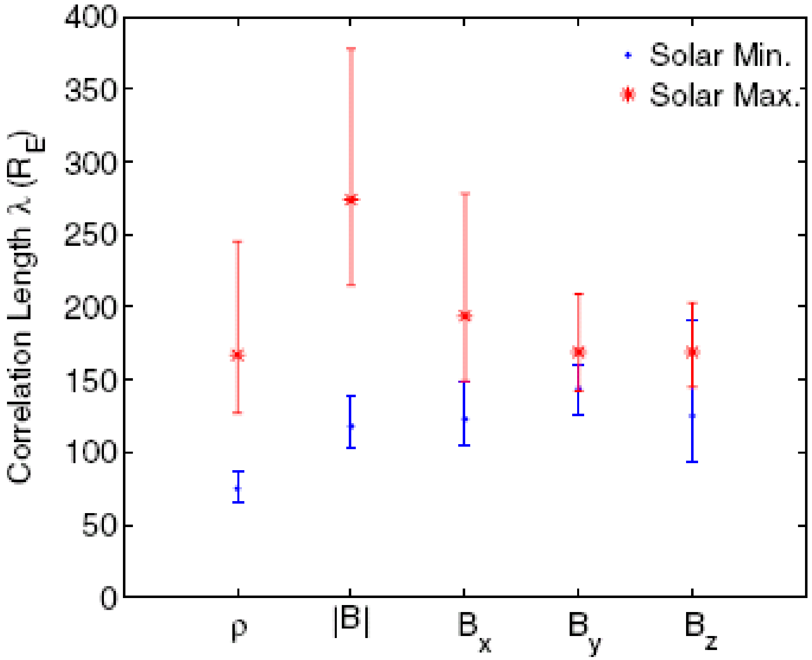
Correlation lengths λ(|B|) and λ(|ρ|) are twice as big at solar max than solar min, and vary together throughout the solar cycle

Correlation lengths λ(B_x, B_y, B_z) do not change with solar cycle

These results may reflect different MHD characteristics (shear vs. compressional Alfvénic) of the nonlinear structures; or at larger scales, differences between structure synthesis in the solar corona and evolution in the solar wind

Linear cross-correlation yields similar trends

Decay lengthscale of linear cross-correlation λ in units of Earth radii R_E



To compute cross-correlation, we need averages defined on a time window with duration τ:

- Short τ ~ 200 minutes enables us to resolve MHD turbulent fluctuations in the inertial range
- Long τ ~ 960 minutes averages over these, so that we focus on the 1/f range of propagating large scale fluctuations that may retain information about their coronal origin

We find solar cycle dependence of correlation measures is independent of window size

CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe

Conclusions

- Information is a physical quantity
- The flow of information is central to the generic physics of complex systems, notably including natural and fusion plasmas
- Understanding the flow of information provides unique, and uniquely satisfying*, pathways to physical understanding of the operation of the system
- Quantitative measures of information are available, which
  - work for real plasma data
  - complement other measures of correlation

*partly because it is the closest we physicists ever get to causation

Acknowledgments

Sandra Chapman, Thomas March, Robert Wicks

CFSA centre for fusion, space and astrophysics

UKAEA Fusion Working in Europe