# Information Theory and Plasma Turbulence

R. O. Dendy[a,b]

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

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

**Abstract.** Information theory, applied directly to measured signals, yields new perspectives on, and quantitative knowledge of, the physics of strongly nonlinear and turbulent phenomena in plasmas. It represents a new and productive element of the topical research programmes that use modern techniques to characterise strongly nonlinear signals from plasmas, and that address global plasma behaviour from a complex systems perspective. We here review some pioneering studies of mutual information in solar wind and magnetospheric plasmas, using techniques tested on standard complex systems.

**Keywords:** information theory, plasma turbulence, complex systems.
**PACS:** 87.19.1o, 89.75.-k, 52.35.Ra

## 1. Introduction

Information theory[1] offers a fresh and productive perspective on plasma turbulence. At the technical level, it offers new ways of quantifying nonlinear behaviour. At the phenomenological and conceptual level, it provides new insights, reflecting the fact that information is a physical quantity. Plasmas support diverse collective nonlinear processes, operating over an exceptionally wide range of lengthscales and timescales – spanning respectively factors of $10^4$ and $10^{10}$ in fusion experiments, for example – whose coupled interactions give rise to the observed emergent phenomenology. The need to interpret, predict, and control this, both for fusion and for solar-terrestrial applications, motivates the present paper.

In a given small volume of plasma, the distribution of charge, current, and electromagnetic field at any moment represents a self-consistent solution of the coupled nonlinear equations governing the dynamics of multiple charged particles together with Maxwell's equations. Given the very large number of particles involved, direct mathematical solution is impossible and physically motivated truncation – the construction of reduced equations – is called for[2]. Thus phenomenology within specific intermediate ranges of lengthscales and timescales is addressed mathematically using different levels of description obtained by truncating, or averaging over, the fullest kinetic description. For example, averaging over gyro-angles yields the gyrokinetic description, while averaging over lengthscales long compared to gyro-orbits yields the magnetohydrodynamic (MHD) fluid description. However, phenomena that are primarily describable using a particular level of model are often strongly coupled to those operating at another level. For example, the stability of quasi-fluid

CP1188, *2009 ICTP Summer College on Plasma Physics*
edited by Bengt Eliasson and Padma K. Shukla

MHD modes of oscillation can be critically affected by kinetic resonant wave-particle interactions. Importantly also, each averaging process and mathematical truncation must result in loss of information.

Turning from the equations to their boundary conditions, we note that fusion and solar-terrestrial plasmas are open systems: energy and information are continuously transferred across their boundaries. Information bears on causation and prediction, and we note that plasmas are far from completely predictable: unexpected behaviour can emerge – prediction becomes expectation. For example, plasma performance in future magnetically confined fusion 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 [3] and edge localised modes [4] in tokamaks, were discovered 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 apparent 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 heating and fuelling. In both cases, history is crucial and there is an element of irreversibility, implying a link to entropy and thence to information.

Returning to basic plasma characteristics, we note that their typical very low density and high temperature imply 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. On the one hand, particle-in-cell codes which implement this lowest-level description (albeit already in terms of representative averaged 'macroparticles') are best adapted to phenomena occurring on the fastest timescales and shortest lengthscales. On the other hand, as noted, any higher level description constitutes a reduced model, in whose construction, information is deliberately eliminated.

How then are we to evaluate the information contained in plasma systems, and what can we learn therefrom? This question is of interest beyond plasma physics, because it appears that information theory may in future provide unifying principles for complex systems science in general. Irrespective of their physical and mathematical embodiment, all complex systems have in common the creation, transmission, sharing and destruction of information. It is this 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, such as a plasma, is thus crucial both to understanding its working, and to rigorously characterising its behaviour. We outline below some pioneering studies of mutual information[5-7] in solar wind and magnetospheric plasmas, using techniques tested [8] *inter alia* on a standard complex systems model [9] for the collective dynamics (i.e., flocking) of birds. While significant progress is being made, 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.

In this paper we focus on applications of information theory to correlated spatiotemporally separated measurements of turbulence in the solar wind plasma. Sections 2 and 3 provide contextual material.

## 2. What is information?

Information resides in the number of yes/no (equivalent to binary 0/1) questions (equivalently bits) to which we have the answer. For example, for $n = 3$ questions there are: $2^3 = 8$ possible combinations of yes/no answers, expressible as 8 three-bit binary symbols 101, 110,…, etc. In general, to represent $n$ bits requires an alphabet containing $M = 2n$ symbols. Suppose we sample (at each sampling we 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:

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

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 $X_1$, $X_2$, ..,$X_i$,… ,$X_N$ drawn from an alphabet having $M = 2n$ symbols. Different symbols $X_i$ recur $N_i$ times, implying different empirical probabilities $P_i = N_i/N \neq 1/M$. Intuitively, the occurrence in the signal of a statistically rare symbol (small value of $P_i$, e.g. letter “x” in English) 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 that the information per symbol is $H/N = - \log_2 P$. It therefore appears logical to define the information gained from a single occurrence of $X_i$ as -$\log_2 P_i$.

Suppose that in a signal of length $N$ symbols, $X_i$ occurs $N_i$ times. Then the total information provided by the occurrences of $X_i$ is $H_i = - N_i \log_2 P_i$, and the total information in the signal is accordingly

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

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, so called because of deep analogies with statistical mechanical and thermodynamic entropy.

Suppose we have two signals – that is, time series of data measurements – labelled A and B. These can be partitioned into an alphabet, meaning a discrete set which spans all the values which the signal can take. Then each measurement within A is a member of the set $\{a_1, \ldots a_i, \ldots a_n\}$, where $a_1$ and $a_n$ are the minimum and maximum values that A is found to take. Within the discretised signal A, each value $a_i$ is found to occur with a probability $P(a_i)$, and similarly for each element $b_i$ in B, we find $P(b_i)$. For the joint probability of $a_i$ in A and $b_j$ in B we have $P(a_i,b_j)$. As we have discussed, the rarer an element $a_i$, the more information its occurrence conveys, an observation that underlies Shannon's definition[1] $H(\mathrm{A}) = - \Sigma\, P(a_i)\log_2 P(a_i)$ of the entropy of a signal analysed in this way. The mutual information $I(\mathrm{A,B})$ between two signals is then $H(\mathrm{A}) + H(\mathrm{B}) - H(\mathrm{A,B})$, and the normalised mutual information is

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

Information is both a physical quantity and an intrinsically nonlinear measure of correlation. In these respects it differs fundamentally from the conventional linear cross-covariance

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

Plasma systems often yield highly nonlinear measurements that are intermittent and 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.

As an example, Fig.1 shows both the mutual information and the classical susceptibility calculated for an implementation of the model from Ref.[9] of flocking birds. In this model, each flying bird takes account of the velocity orientation of its near neighbours, and does its best (subject to noise) to align with them. Speed is constant, but 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, such that

$$\begin{aligned} x_{n+1} &= x_n + \vec{v}\,\delta t \\ \theta_{n+1} &= \langle\theta_n\rangle_R + \delta\theta_n \end{aligned}$$

where the noise range is $-\eta < \delta\theta < \eta$.

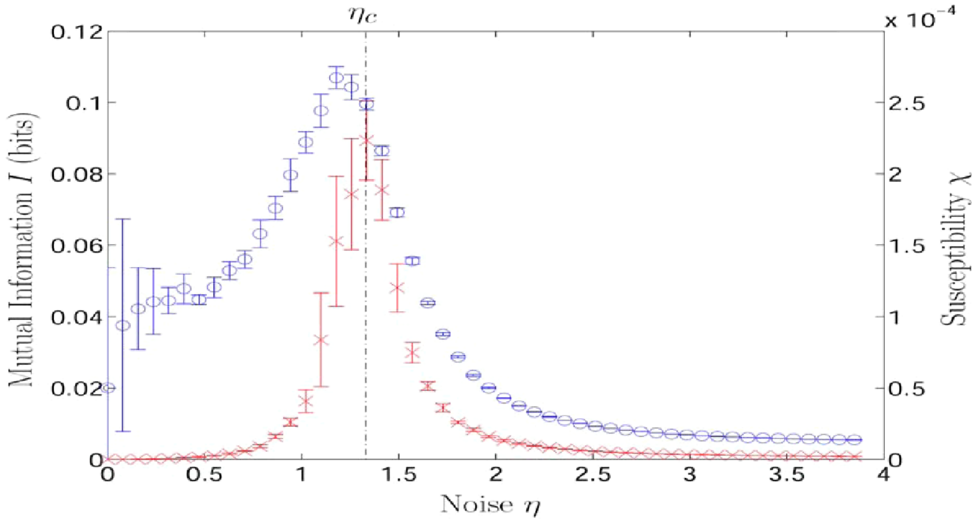


Fig.1 Measured mutual information (blue) and susceptibility (red) versus noise η for the model of Vicsek *et al.*[9] for bird flocking. From Ref.[8].

It is evident from Fig.1 that the information theoretic measure yields a perspective that is complementary to the classical one. For example, error bars near the peak identifying the phase transition are at their smallest for mutual information, but at their largest for susceptibility.

As we shall discuss in Sections 4 and 5, quantifying the mutual Shannon information shared between two causally linked but spatiotemporally separated plasma signals can identify key timescales, distinguish between plasma physics models for the propagation of perturbations, and measure the strength of the causal link. Information theory offers a fresh perspective[5-7] on nonlinear phenomenology and turbulence in plasmas.

**3. Solar wind plasma turbulence: a brief introduction**

The solar wind is a supersonic plasma flow which originates from the solar corona and propagates through interplanetary space, filling it until it reaches the local interstellar medium at the heliopause. Solar wind fluctuations include both large scale episodic nonlinear perturbations and turbulence. The solar wind provides unique opportunities for long duration *in situ* studies of magnetohydrodynamic turbulence in a plasma flowing supersonically with high magnetic Reynolds number ~$10^5$. Its spectral power density scales approximately as inverse frequency $f^{-1}$ at lower frequencies (≤1 mHz); and as $f^{-5/3}$, reminiscent of Kolmogorov's inertial range, at higher frequencies (~10–100 mHz). Both the $f^{-5/3}$ and $f^{-1}$ fluctuations are often predominantly shear Alfvénic in character. The frequency at which the transition between power laws occurs (~1–10 mHz) is observed to decline with increasing heliocentric distance in the plane of the ecliptic, and this extension of the $f^{-5/3}$ range at greater distances can be interpreted as evidence for an evolving turbulent cascade. The $f^{-1}$ range is taken to reflect embedded solar coronal turbulence, convected with the solar wind, while the large-scale magnetic structure of the corona varies with the solar cycle and heliospheric latitude, creating variations in solar wind speed.

The solar wind thus provides a natural laboratory for studying strongly nonlinear and turbulent phenomenology in a magnetised plasma that shows, for example, a clear inertial range on timescales from minutes to hours, and whose magnetic Reynolds number is of order $10^5$. Spacecraft located in the solar wind upstream of Earth's bow shock have the further advantage of sampling this high Mach number turbulent plasma flow far from any boundary layer. The signals obtained are often strongly nonlinear in character and are sparse, in the sense that observations are available at only a very small number of points in space, depending on the number of spacecraft that are simultaneously taking measurements of related plasma parameters in the upstream solar wind. In order to characterise the observed signals, and relate them to plasma models, it is therefore necessary to quantify spatial correlation within a strongly nonlinear system for which information is sparse. This can be achieved both by conventional analysis techniques such as generalised structure functions and extended self similarity (for recent applications to the solar wind, see for example Ref.[10]), and through information theory, as we now describe.

## 4. Information theory applied to causally linked signals measured in the upstream solar wind and in the terrestrial ionosphere

Studies of spatiotemporal correlation between coupled nonlinear signals in the solar wind and magnetosphere-ionosphere, reported in Refs.[5,6], show that the mutual Shannon information, and associated recurrence plot techniques, are helpful in this context. We refer to these papers for a detailed account of how physically relevant timescales are extracted from these studies. In outline, Fig.2 shows measured fluctuations in solar wind energy from the WIND satellite at the sunward libration point, together with contemporaneous terrestrial magnetometer data (AE – auroral electrojet) at high geomagnetic latitude. These are causally linked because the solar wind drives magnetotail reconnection; following a succession of coupled nonlinear plasma processes occurring on different lengthscales and timescales within the magnetosphere, this energy release drives ionospheric currents which affect the terrestrial magnetic field, as measured. Clearly the signals in Fig.2 share features in common, which appear to be shifted in time relative to each other.

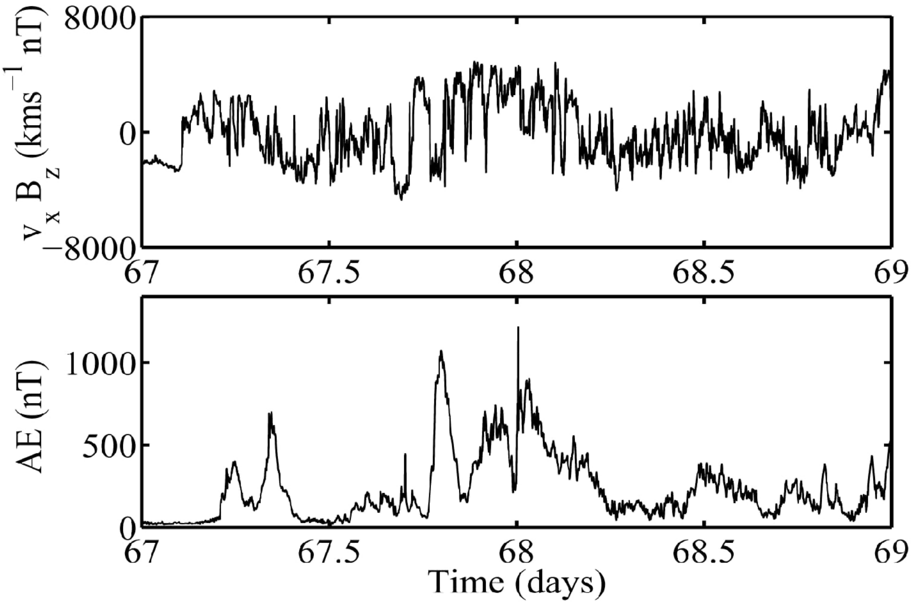


Fig.2 Measurements of solar wind energy fluctuations (upper) from the WIND satellite and terrestrial magnetometer data (lower) during the first half of 1995 around solar minimum. For details and acknowledgments see Refs.[5,6].

This causal linkage can be quantified in terms of the Shannon information shared between the two signals, in a way which sheds light on the underlying plasma physics. Specifically, the propagation time $\Delta t$ that one infers for a fluctuation in the solar wind travelling from the WIND satellite to the edge of the magnetosphere, will depend on the model chosen for solar wind structure. Given such a model (several are available, and are used in Ref.[6]), $\Delta t$ is calculated from the spacecraft location combined the measured velocity **v** and local mean magnetic field **B** of the solar wind. There remains a second unknown time lag $\Delta t'$ arising from the sequence of plasma processes that transmit the effect of the incident fluctuation from the edge of the magnetosphere down to the ionosphere.

Given the physically understood causal linkage, a key question to ask is: what combined timeshift $\Delta t + \Delta t'$ maximises the mutual information between the two signals? This is answered in Fig.3. This quantifies shared information as a function both of the solar wind propagation models driving the value of $\Delta t$ and of the unknown time lag $\Delta t'$ which is treated as a free parameter. We infer from Fig.3 that, first, the Parker spiral model is most effective and, second,

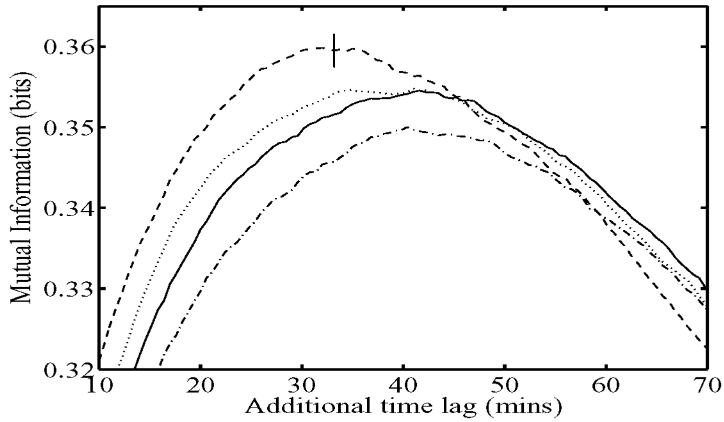


Fig.3 Mutual information shared between the signals in Fig.2 as a function of time lag $\Delta t'$ introduced by plasma processes within the magnetosphere. The different curves realte to different models for propagation of solar wind fluctuautions from the libration point to the edge of the magnetosphere, giving rise to different values of $\Delta t$. See Ref.[6].

that the sequence of plasma processes within the magnetosphere take about 30 minutes. Thus by treating the physical problem as an information theoretic problem of signal propagation (from the libration point to the edge of the magnetosphere) and signal processing (within the magnetosphere), paradoxically we extract fresh physical information. We refer also to Ref.[5] for details of the visualisation and quantification of information in this context using recurrence plots.

## 5. Information theory applied to spatiotemporally separated correlated signals measured in the upstream solar wind

As a second example, let us consider a recent information theoretic analysis[7] of data obtained during periods from 1998 onwards when the Wind, ACE and Cluster spacecraft were simultaneously in the upstream solar wind, and explored a range of spatial scales sufficient to determine correlation properties[11]. Nonlinear correlation is quantified here by calculating the mutual information between measurements, for example those in Fig.4, of magnetic field $\boldsymbol{B}$ (magnitude and individual vector components), flow velocity $\boldsymbol{v}$, and density $\rho$ from spatially separated spacecraft. This enables us to compare the relative degree of correlation between different solar wind bulk parameters. The ordering of mutual information with respect to signal propagation relative to the background magnetic field direction is then related to models and understanding of anisotropic solar wind plasma turbulence.

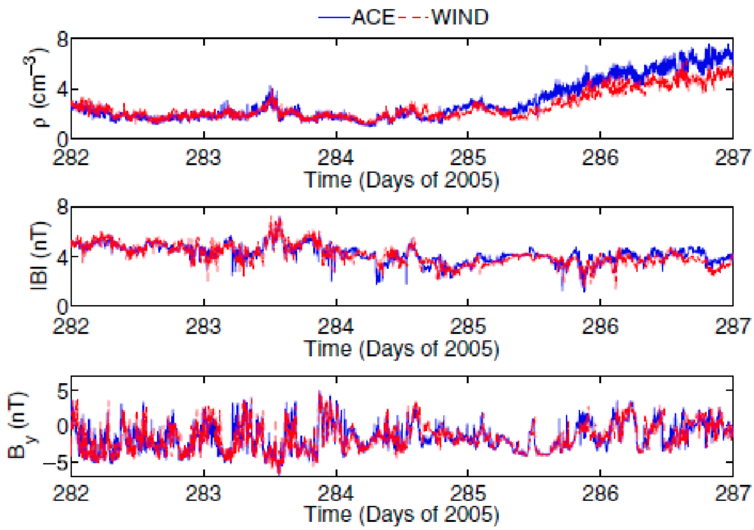


Fig.4 A typical five day period of solar wind observations from the *ACE* (solid blue line) and *WIND* (dashed red line) spacecraft. The top panel shows the ion density $\rho$ ($cm^{-3}$), the middle panel the magnetic field magnitude $|B|$ (nT), and the bottom panel the $B_y$ (nT) component of the magnetic field. From Ref.[7].

Figure 5 shows plots[7] of normalised mutual information versus spatial separation between the ACE and WIND spacecraft, which took contemporaneous measurements in the upstream solar wind. These measurements were taken under a range of solar conditions, and it is clear from Fig.5 that the spatial decline of shared information provides an effective measure of the extent of correlation of magnetic field strength, magnetic field components, and plasma density, and of the variation of this correlation with solar activity. This yields physical information, for example differentiating between the pairs of variables $B$ and $\rho$, which would be associated with compressional Alfvénic structures in the linear limit, and $B_x$ and $B_y$, which would be associated with shear Alfvénic structures. The correlation length of the former varies with solar cycle, but not the latter.

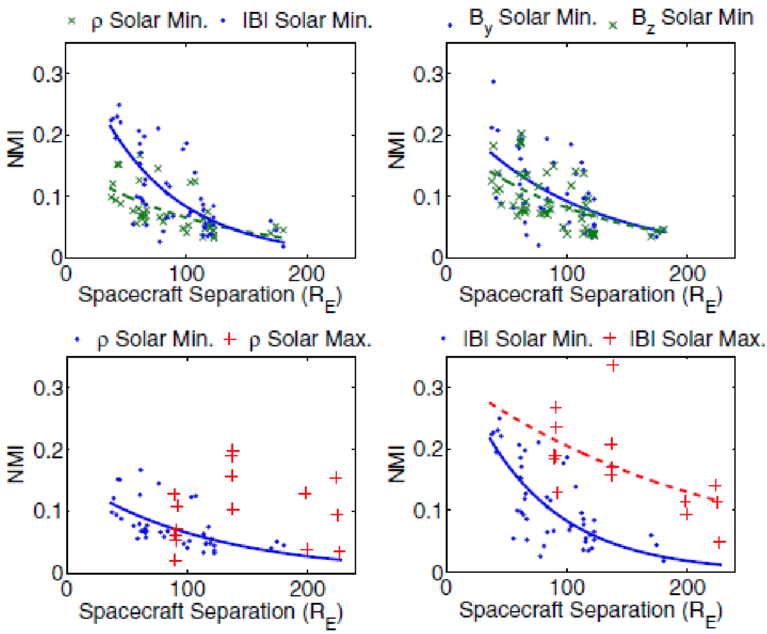


Fig.5 Normalised mutual information between contemporaneous measurements of fluctuations in the solar wind by the ACE and WIND spacecraft, as a function of spacecraft separation in units of the Earth-Sun distance $R_E$, during periods of minimum and maximum solar activity. From Ref.[7].

## 6. Conclusions

Information theory, applied directly to measured signals, yields new perspectives on, and quantitative knowledge of, the physics of strongly nonlinear and turbulent phenomena in plasmas. It represents a new and productive element of the topical research programmes that use modern techniques to characterise strongly nonlinear signals from plasmas (for a review, see Ref.[12]), and that address global plasma behaviour from a complex systems perspective (for a review, see Ref.[13]). The successful examples of information theory applied to the solar wind and magnetosphere motivate future studies in other plasma contexts.

**Acknowledgments** It is a pleasure to acknowledge the contributions of several collaborators on this topic at Warwick, including Sandra Chapman, Thomas March, Ruth Nicol, and Robert Wicks. This work was supported in part by Euratom (but the views and opinions expressed herein do not necessarily reflect those of the European Commission) and by the United Kingdom Engineering and Physical Sciences Research Council.

## References

[1] Shannon C E and Weaver W, *The Mathematical Theory of Communication* (Champaign: University of Illinois Press, 1949)
[2] For a brief accessible introduction to the fundamentals of plasma physics, see for example Dendy R O, *Plasma Dynamics* (Oxford: Oxford University Press, 1990)
[3] ITER Physics Expert Group on Confinement and Transport *et al.*, *Nucl. Fusion* **39** 2175 (1999).
[4] Kamiya K *et al.*, *Phys. Control. Fusion* **49** S43 (2007).
[5] March T K, Chapman S C and Dendy R O, *Physica D* **200** 171 (2005).
[6] March T K, Chapman S C and Dendy R O, *Geophys. Res. Lett.* **32** L04101 (2005).
[7] Wicks R T, Chapman S C and Dendy R O, *Astrophys. J* **690** 734 (2009).
[8] Wicks R T, Chapman S C and Dendy R O, *Phys. Rev. E* **75** 051125 (2007).
[9] Vicsek T, Czirok A, Ben-Jacob E, Cohen I and Shochet O, *Phys. Rev. Lett.* **75** 1226 (1995).
[10] Nicol R M, Chapman S C and Dendy R O, *Astrophys. J.* **679** 862 (2008).
[11] Matthaeus W H, Dasso S, Weygand J M, Milano L J, Smith C W and Kivelson M G, *Phys. Rev. Lett.* **95** 231101 (2005).
[12] Dendy R O and Chapman S C, *Plasma Phys. Control. Fusion* **48** B313 (2006).
[13] Dendy R O, Chapman S C and Paczuski M, *Plasma Phys. Control. Fusion* **49,** A95 (2007)