Consiglio Nazionale delle Ricerche — Istituto dei Sistemi Complessi — Sezione di Firenze

# Information Theory and Space Weather


**Massimo Materassi**

massimo.materassi@fi.isc.cnr.it

Istituto dei Sistemi Complessi del Consiglio Nazionale delle Ricerche ISC-CNR, Sezione Territoriale di Sesto Fiorentino (Italy)


Spineto (Italy), June 09, 2012

COST ES0803

1. What is information theory;

2. Information theory and coupled processes: mutual information, delayed mutual information, transfer entropy;

3. Information theory vs Space Weather: cascades, near-Earth plasma physics and ionospheric scintillation, storms and substorms;

4. Conclusions and future work...

# 1. What is information theory;

**Ignorance** is the subject of Information Theory (just like "absolute" laws are the goal of Relativity...)

**Deterministic dynamics**: once the initial conditions are known, the evolution is perfectly known forever!

**Stochastic dynamics**: processes the outcome of which is known only in a probabilistic sense. At each time many outcomes are possible, with different probability to take place.

$$\dot{y} = f\left(y, \xi\right)$$

**Probabilistic terms** in the DEs act as the cylinder hat of a magician.

**Ignorance about a certain process that may have *R* different (equally probable) outcomes:**

$$I = I(R), \qquad \frac{dI}{dR} > 0$$

**Additivity of the ignorance about concurring independent processes:**

$$I(Y_1(R_1), Y_2(R_2)) \equiv I(R_1 R_2) = I(R_1) + I(R_2)$$

**The logarithm of *R* is "the" solution:**

$$I(R) = K \ln R$$

**Ignorance about the outcome of a process with known PDF:**

$$I[p] = -K \sum_y p(y) \ln p(y)$$

$$I(R) = K\ln R = \frac{R}{R}K\ln R = pRK\ln\frac{1}{p} = -RpK\ln p =$$

$$= -\sum_{i=1}^{R} pK\ln p = -K\sum_{i=1}^{R} p_i \ln p_i .$$

**Ignorance of the outcome is information content:**

$$I_i = K\ln R_i = K\ln R, \qquad I_f = K\ln R_f = K\ln 1 = 0$$

$$-\Delta I = I_i - I_f = K\ln R = I(R)$$

**Joint information (ignorance) about any two processes *X* and *Y*:**

$$I(X,Y) = -K \sum_{x,y} p(x,y) \ln p(x,y)$$

**Conditioned information: how ignorant are we about *Y* if we do already know *X*?**

$$I(Y|X) \overset{\text{def}}{=} I(X,Y) - I(X)$$

**The same quantity explicitly written in terms of PDFs:**

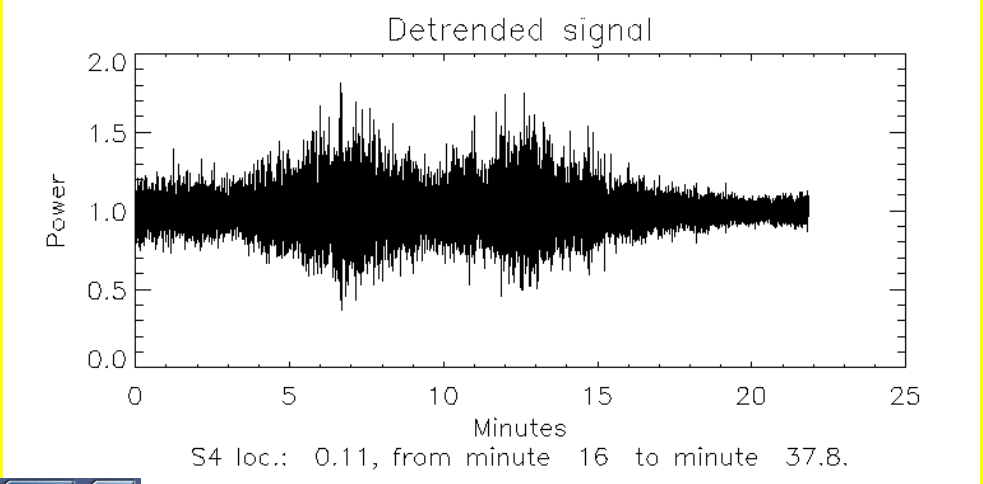
$$I(X|Y) = -K \sum_{x,y} p(x,y) \ln p(y|x)$$

**Need of PDFs: need of inferring ensemble-statistics "from what one has" in Geophysics, i.e. time series.**

**Time-statistics versus ensemble-statistics…**

**Information depends functionally on the PDF of a stochastic process, from which it inherits the same time dependence (non-stationary processes):**

$$p = p(t) \quad \Longrightarrow \quad I[p] = I[p(t)]$$




**Potential *caveat* about the use of time-statistics in the place of ensemble-statistics**

# Ignorance and (thermodynamical) Entropy 1/2

PHYSICAL REVIEW VOLUME 106, NUMBER 4 MAY 15, 1957

## Information Theory and Statistical Mechanics

E. T. JAYNES
*Department of Physics, Stanford University, Stanford, California*
(Received September 4, 1956; revised manuscript received March 4, 1957)

$$\delta I\,[p]\big|_{\text{phys}} = 0$$

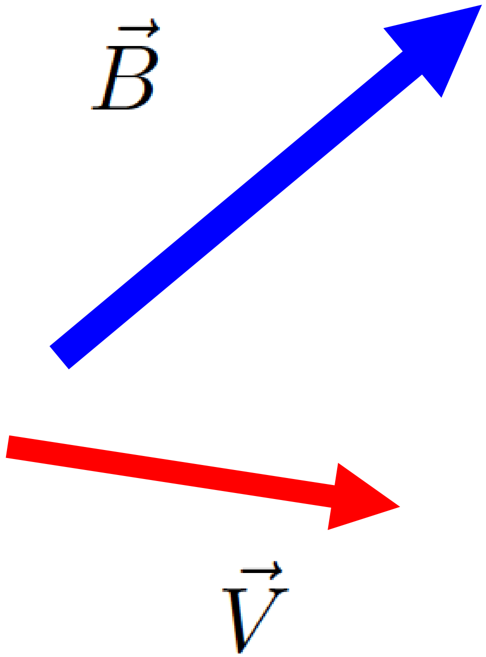
PHYSICAL REVIEW VOLUME 108, NUMBER 2 OCTOBER 15, 1957

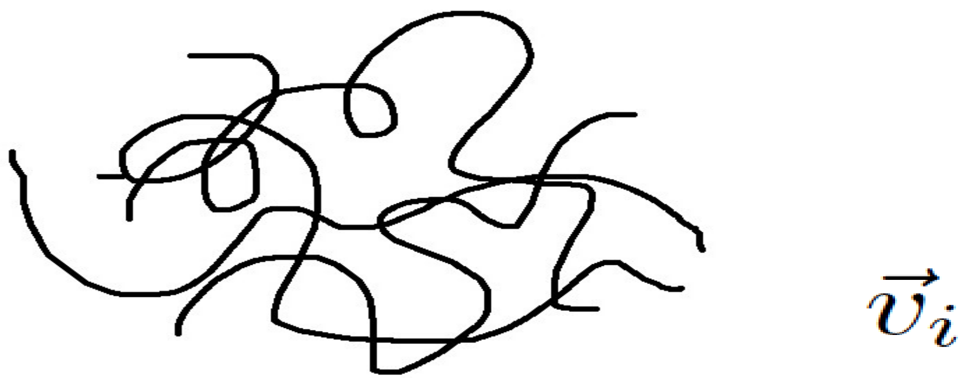
## Information Theory and Statistical Mechanics. II

E. T. JAYNES
*Department of Physics, Stanford University, California*
(Received March 15, 1957)

# Ignorance and (thermodynamical) Entropy 2/2

**Entropy growth due to dissipation: consuming the “deterministic variables” in favour of “stochastic” ones**

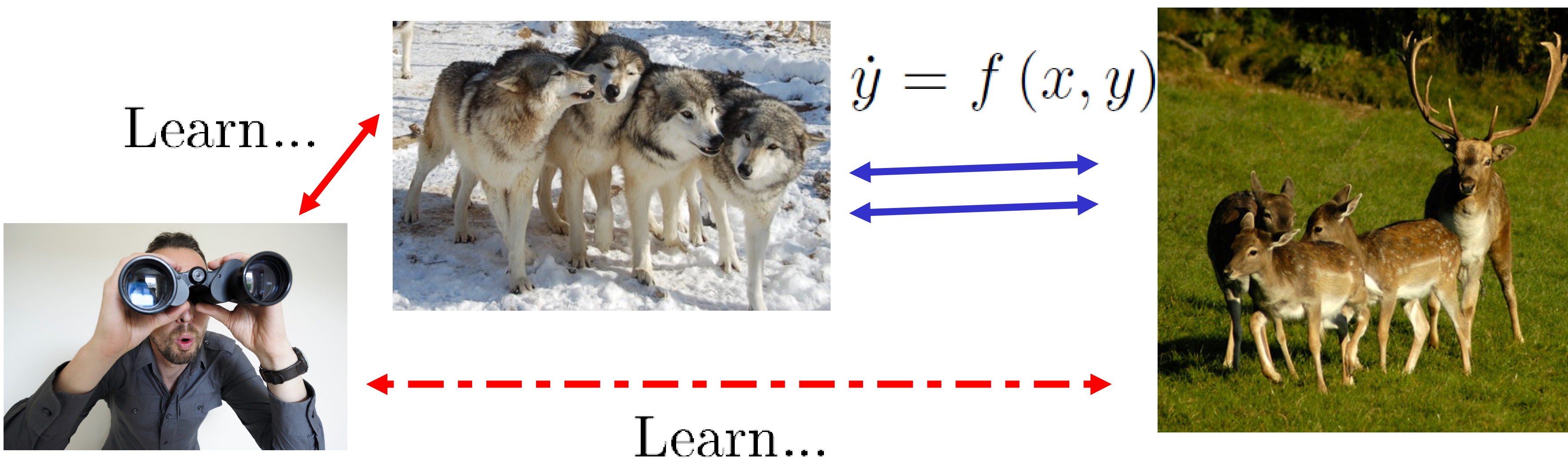


1. What is information theory;

2. Information theory and coupled processes: mutual information, delayed mutual information, transfer entropy;

**Interacting processes may teach something of each other: from interaction (dynamics) to cognition (information theory)…**



**How much information about *Y* can the knowledge of *X* add, if *X* and *Y* are interacting "with no delay"? Mutual information:**

$$M_{Y,X} = I(Y) - I(Y|X)$$

**Mutual information highlights statistical inter-dependence:**

$$M_{Y,X} = K \sum_{x,y} p(x,y) \ln\left(\frac{p(x,y)}{p(x)\,p(y)}\right)$$

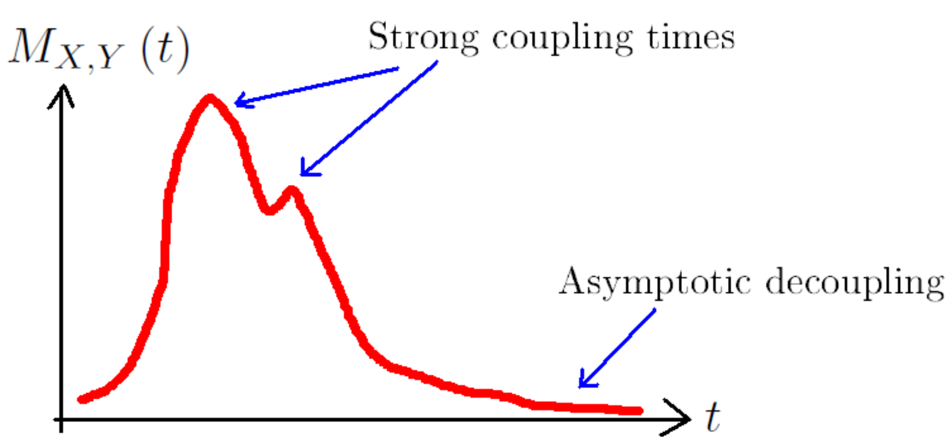
**This involves the PDFs at all orders of momenta, as a fully non-linear correlation.**

**Mutual information is symmetric under *X*⟷*Y*** $M_{Y,X} = M_{X,Y}$

**Mutual information is *the information shared***

**Mutual information (MI) depends on time due to the dependence that *p* shows.**

**The variability of MI with time indicates that the instantaneous degree of coupling of the two processes *X* and *Y* is varying with time.**




**This kind of time dependence renders "when" *X* and *Y* influence each other.**

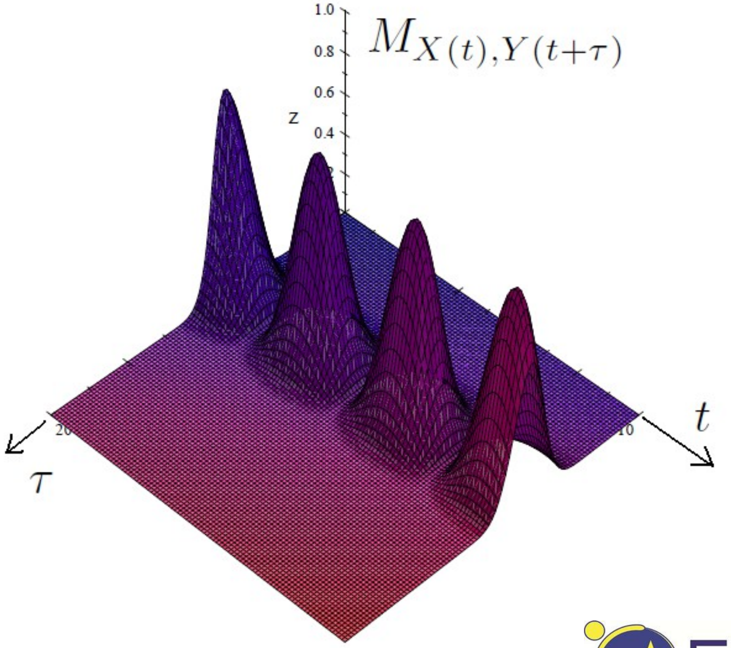
**What if one is interested in the response time?**

**Response time analysis**: how much information about **the future of *Y*** can the knowledge of **the present of *X*** add, if *X* and *Y* are interacting "with **some delay**"? Delayed mutual information:

$$M_{X(t),Y(t+\tau)} \stackrel{\mathrm{def}}{=} I\left(Y\left(t+\tau\right)\right) - I\left(Y\left(t+\tau\right)\left|X\left(t\right)\right.\right)$$

The dependence **on *t*: when** are the processes interacting?

The dependence **on *tau*: with which delay** does the affected process respond?



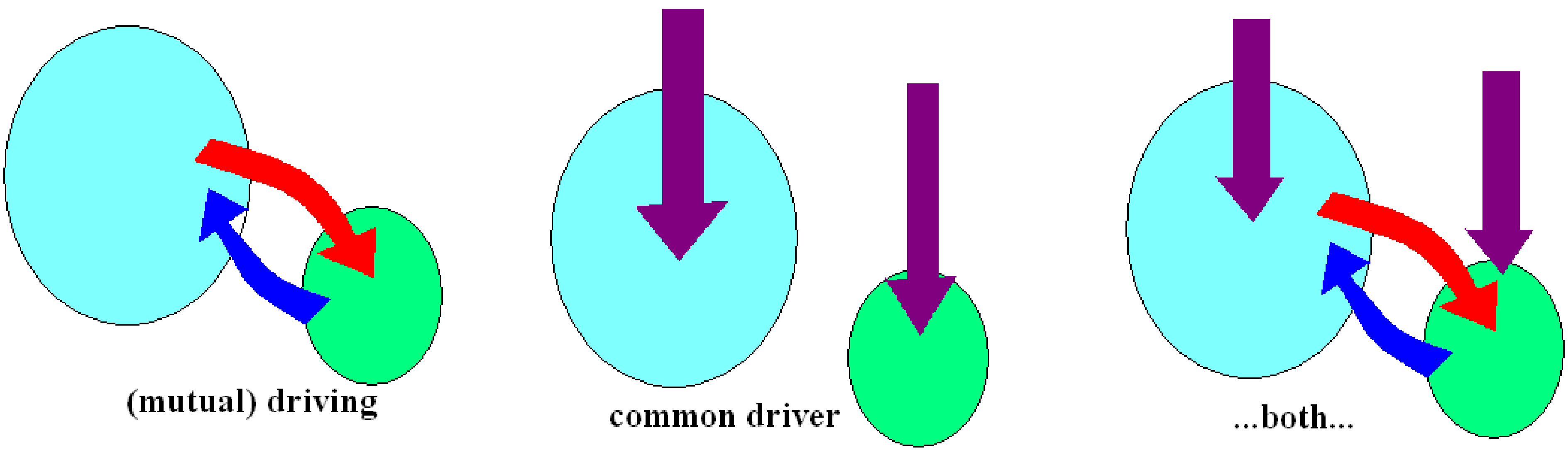
**Master-slave analysis**: which is the process "influencing more" the other one? When? With which "response delay"?

$$\begin{cases} \dot{x} = f\left(x, y\right), \\ \dot{y} = g\left(x, y\right) \end{cases}$$

$$\Delta M_{X(t),Y(t+\tau)} \stackrel{\mathrm{def}}{=} M_{X(t),Y(t+\tau)} - M_{Y(t+\tau),X(t)}$$

The sign of this **differential delayed mutual information** (DMI) indicates which is the process containing more information about the future of the other one, i.e. **which is the process "driving" the other one**

# Interactions and driving: two processes with mutual information can show:



**Transfer entropy: encoding the dynamics of the processes in the DMI using transition probabilities instead of PDFs, or (equally) conditioning the quantities defining the DMI of *X* on *Y* conditioning everything on the present status of *Y* itself:**

$$T_{X\to Y}(t,\tau) \stackrel{\text{def}}{=} I\left(Y(t+\tau)\,|Y(t)\right) - I\left(Y(t+\tau)\,|X(t),Y(t)\right)$$

$$T_{X\to Y}(t,\tau) =$$

$$= K\sum\nolimits_{y(t+\tau),x(t),y(t)} p\left(y(t+\tau),x(t),y(t)\right)\ln\left(\frac{p\left(y(t+\tau),x(t),y(t)\right)p\left(y(t)\right)}{p\left(y(t+\tau),y(t)\right)p\left(x(t),y(t)\right)}\right)$$

**In the transfer entropy contributions from "common drivers" are expected not to appear. Differential transfer entropy:**

$$\Delta T_{X\to Y}(t,\tau) = T_{X\to Y}(t,\tau) - T_{Y\to X}(t,\tau)$$

**Comparison between delayed mutual information and transfer entropy. When all processes are Gaussian, then transition probabilities are calculated in terms of correlation matrices (Kaiser A. & Schreiber T. 2002. Information transfer in continuous processes. Physica D 166: 43–62)**

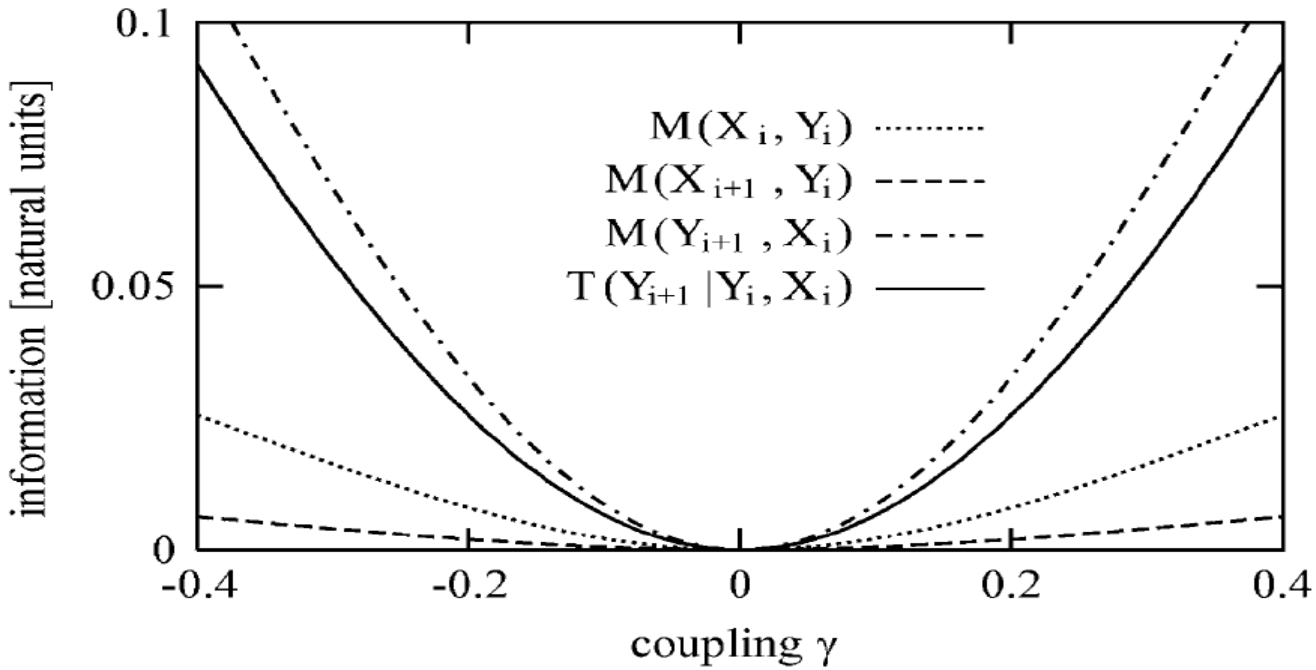
$$M(\mathbf{X}_i^{(k)}, \mathbf{Y}_j^{(l)}) = \frac{1}{2}\log\frac{\det C(\mathbf{X}_i^{(k)})\det C(\mathbf{Y}_j^{(l)})}{\det C(\mathbf{X}_i^{(k)} \otimes \mathbf{Y}_j^{(l)})}.$$

$$T(X_{i+1}|\mathbf{X}_i^{(k)}, \mathbf{Y}_j^{(l)}) = \frac{1}{2}\log\frac{\det C(\mathbf{X}_i^{(k)} \otimes \mathbf{Y}_j^{(l)})\det C(X_{i+1} \otimes \mathbf{X}_i^{(k)})}{\det C(X_{i+1} \otimes \mathbf{X}_i^{(k)} \otimes \mathbf{Y}_j^{(l)})\det C(\mathbf{X}_i^{(k)})}$$

$$X_{i+1} = \alpha X_i + \eta_i^X,$$

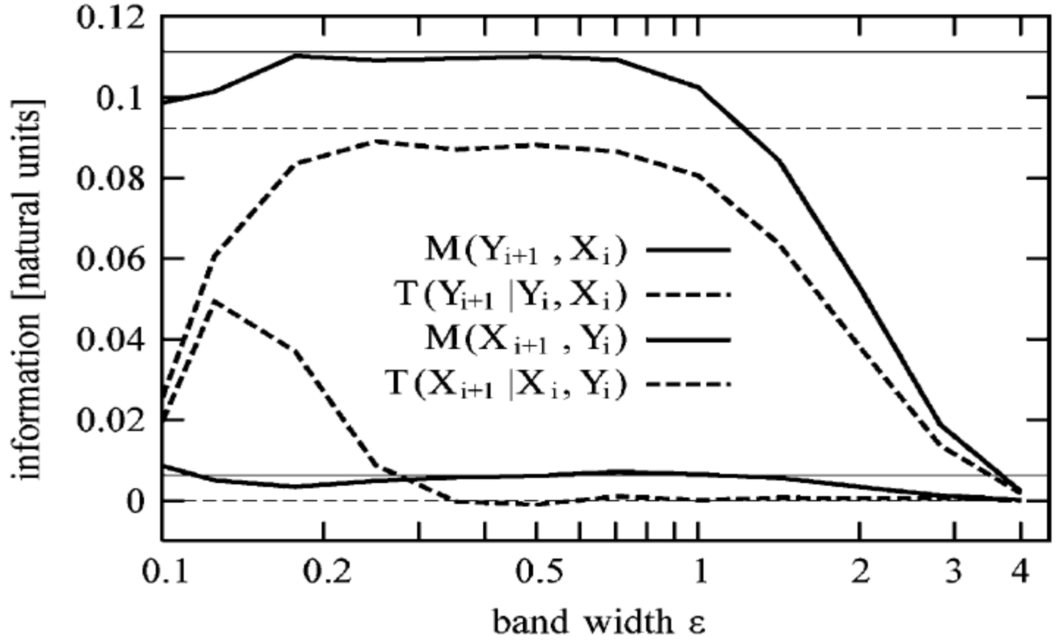
$$Y_{i+1} = \beta Y_i + \gamma X_i + \eta_i^Y.$$

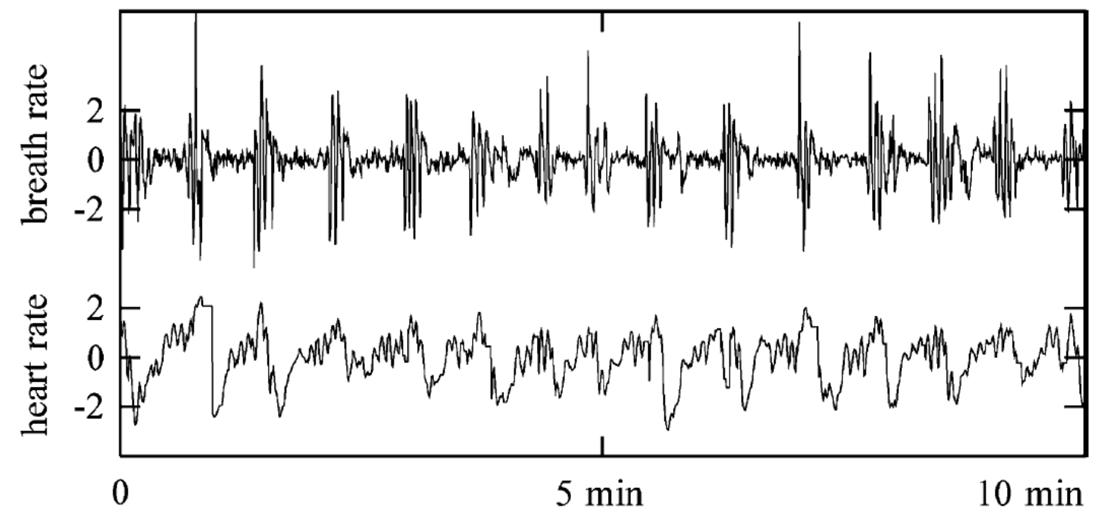
**Example: two linear autoregressive processes coupled and stirred by Gaussian noises**



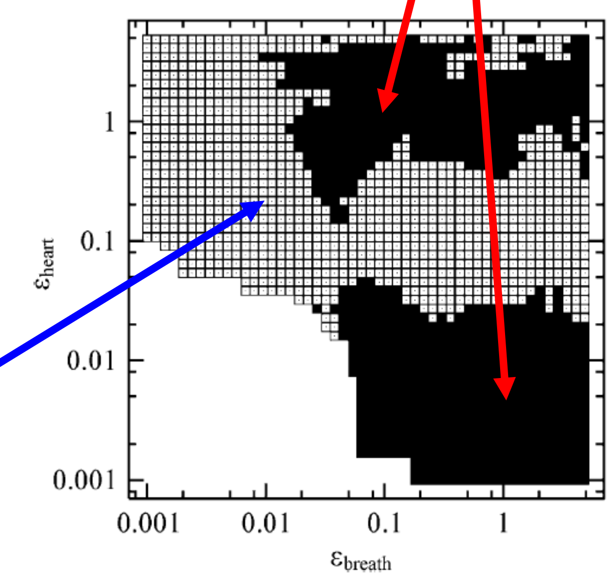
**Binning problem**: in order to build PDFs from time series one will make histograms, that may be dependent on the choice of bins in a critical way, with respect to the calculation of transfer entropy!

A. Kaiser, T. Schreiber / Physica D 166 (2002) 43–62






**Heart drives here**




**Binning problem may render some cases indecidible!**

**Breath drives here**

The **optimal number *M* of bins** for the evaluation of a uniform bin-width histogram is provided by the maximum of the posterior probability (Knuth, 2005):

$$p(M \mid N, n_k) \propto \left(\frac{M}{V}\right)^N \frac{\Gamma(M/2)}{\Gamma(1/2)^M} \frac{\prod_k \Gamma(n_k + 1/2)}{\Gamma(N + M/2)}$$

1. What is information theory;

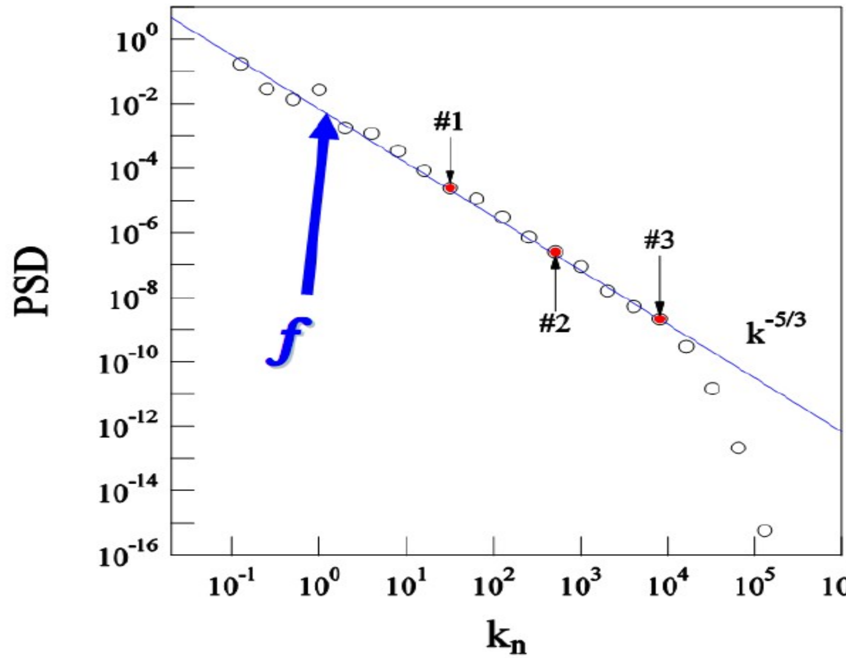
2. Information theory and coupled processes: mutual information, delayed mutual information, transfer entropy;

3. Information theory vs Space Weather: cascades, near-Earth plasma physics and ionospheric scintillation, storms and substorms;

## IT analysis 1/4: Synthetic 1D shell model of direct cascade

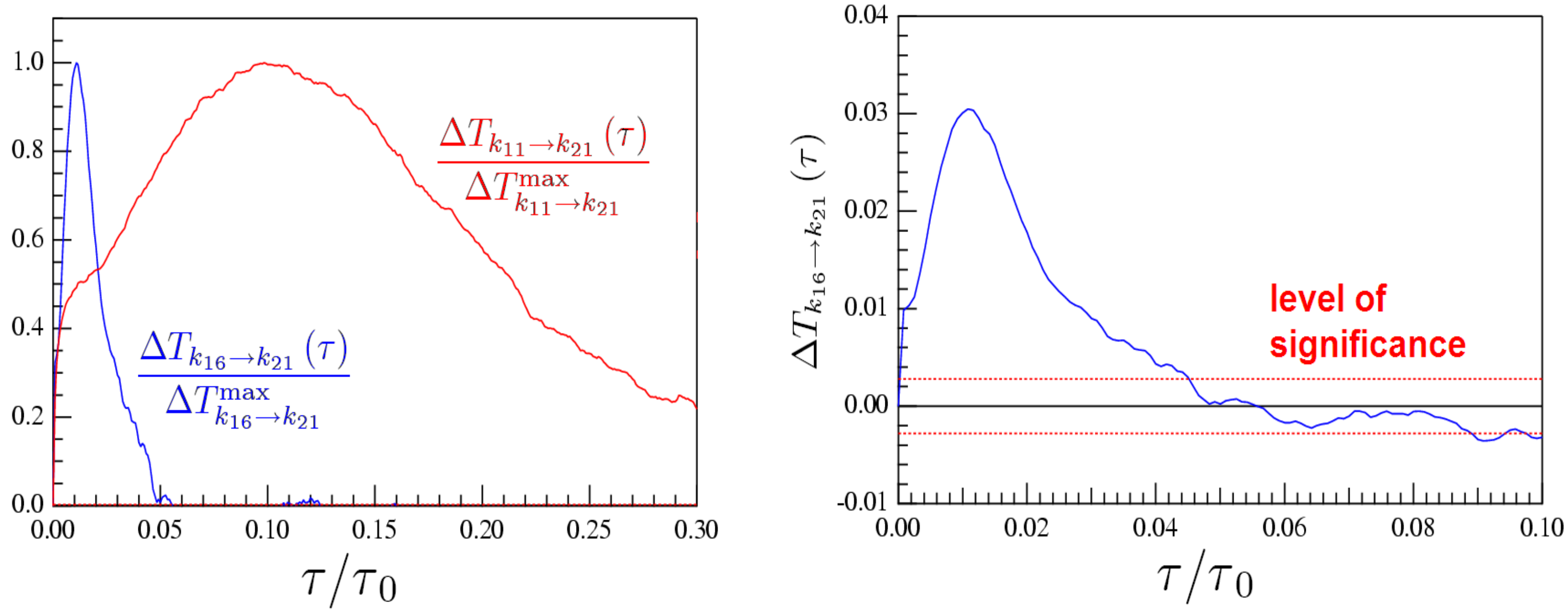
Materassi, M., G. Consolini, N. Smith, in prepatation.

$$\begin{cases} \dot{u}_n = i\left(k_n u_{n+1} u_{n+2} - \frac{1}{2} k_{n-1} u_{n+1} u_{n-1} - \frac{1}{2} k_{n-2} u_{n-1} u_{n-2}\right)^* - \nu k_n^2 u_n + f_n, \\ f_n = \delta_{4,n}\,(1+i)\, f, \end{cases}$$



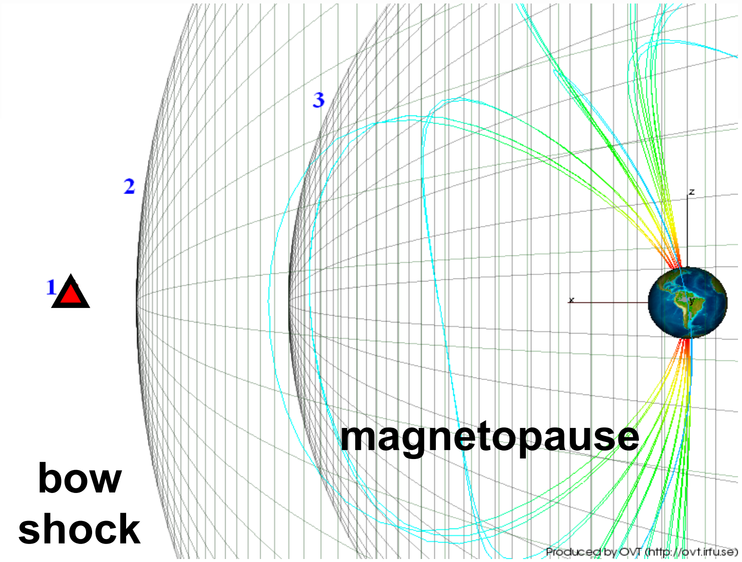

- Coupling in both verses, but prevalence of direct cascade (viscosity suppresses back-reaction, growing with k);
- Injection of energy at a large scale (that pertaining to $n = 4$, here);
- Large scales "drive" small scales locally in the k-space

# The direct cascade and the *k*-local coupling are both clearly visible



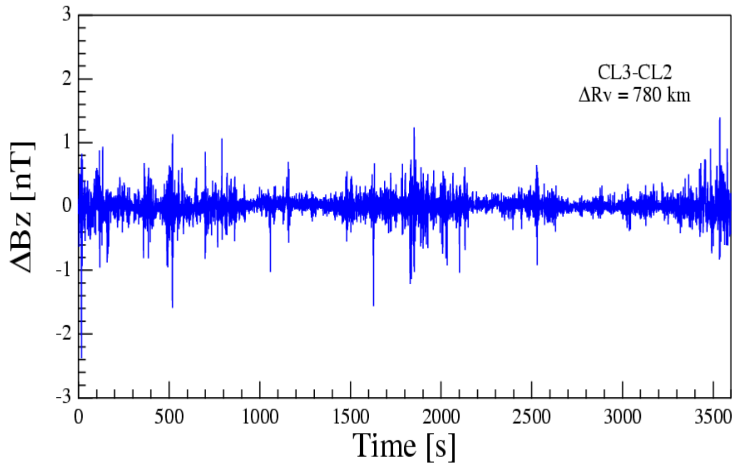
# IT analysis 2/4: Detection of magnetic structure coalescence in real data from Cluster mission

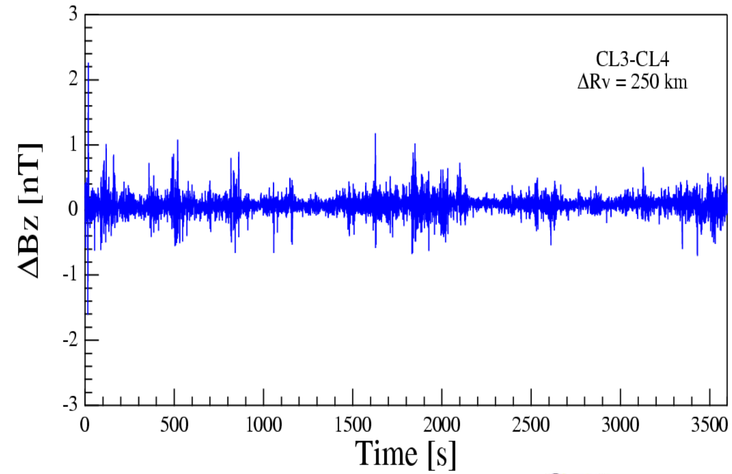
Materassi, M., G. Consolini, N. Smith, in prepatation.



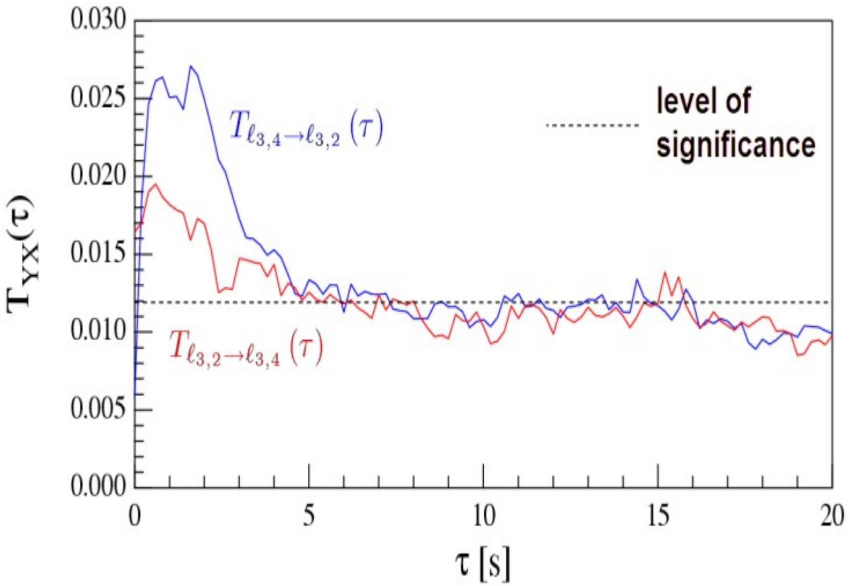

**Cluster data of the z-component of interplanetary magnetic field between 22:30:00 UT and 23:30:00 UT on April 5, 2001 (right outside the bow shock), with a resolution of 5 samples per second**

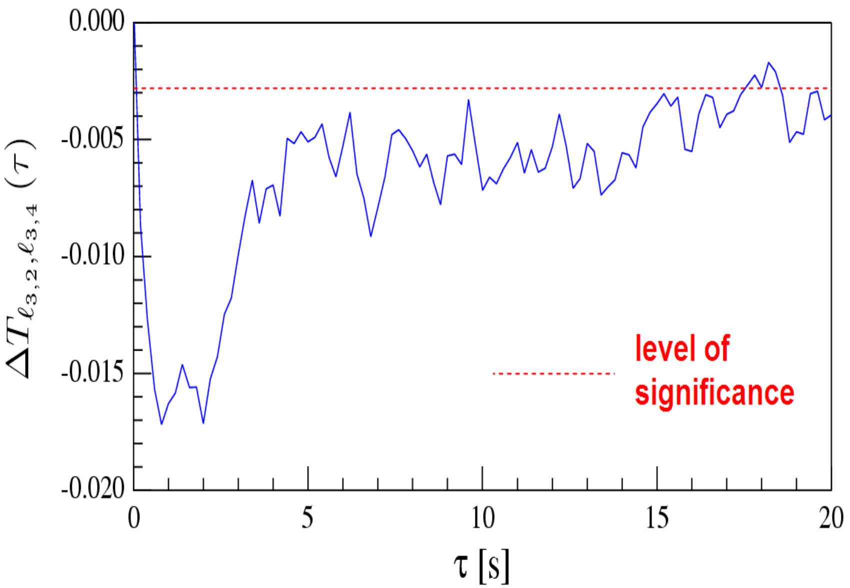
$$\Delta B_z(t,\ell) = B_z(\mathrm{P}_i,t) - B_z\left(\mathrm{P}_j, t+\frac{\ell}{v}\right)$$



## A magnetic structure coalescence (inverse cascade?) appears: structures are probably smashed against the bow shock and coalesce

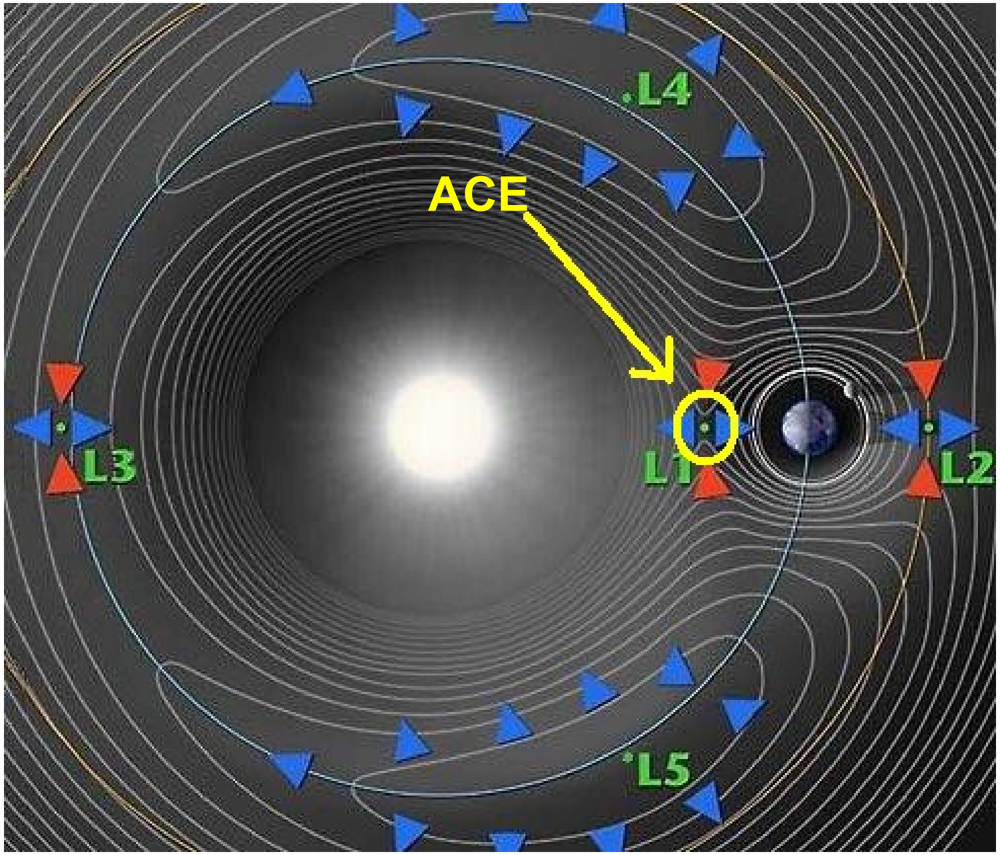


## IT analysis 3/4: relationship between solar wind quantities at ACE location and ionospheric irregularities causing radio scintillation

Please cite this article as: Materassi, M., Ciraolo, L., Consolini, G., Smith, N., Predictive Space Weather: An information theory approach, *Advances in Space Research* (2010), doi: 10.1016/j.asr.2010.10.026

**Link the features of the Sun-Earth forcing with those of the irregular ionosphere**

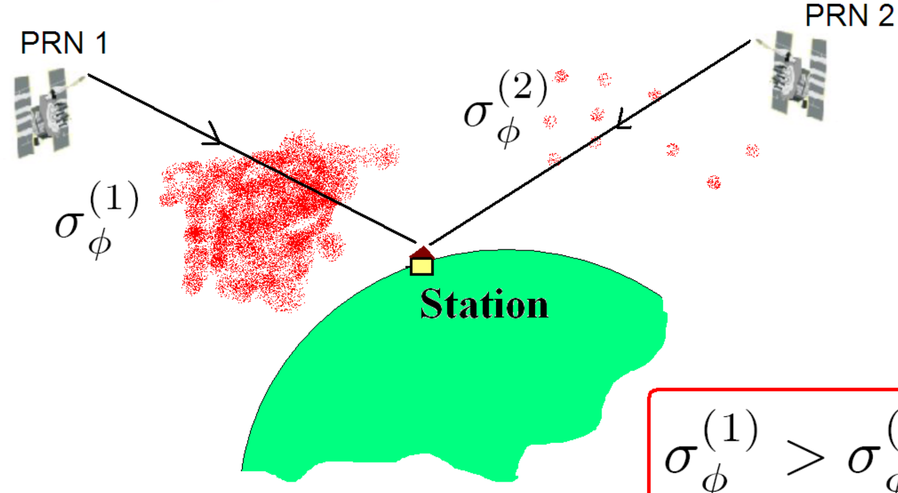



**Putting together Solar-Wind time series…**

$$B_z(t), \quad N_i(t)$$

**…and a time series describing GPS phase scintillation.**

$$\sigma_\phi(t)$$

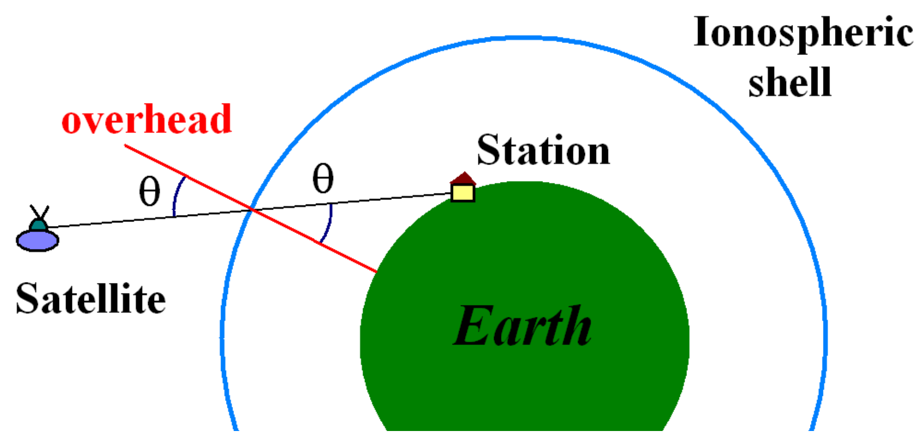



$$\sigma_{\phi}^{(1)} > \sigma_{\phi}^{(2)}$$

**Due to the patchy nature of radio scintillation, scintillation indices do not describe the "state of scintillation" on the top of a location, but just along the radio link raypath.**

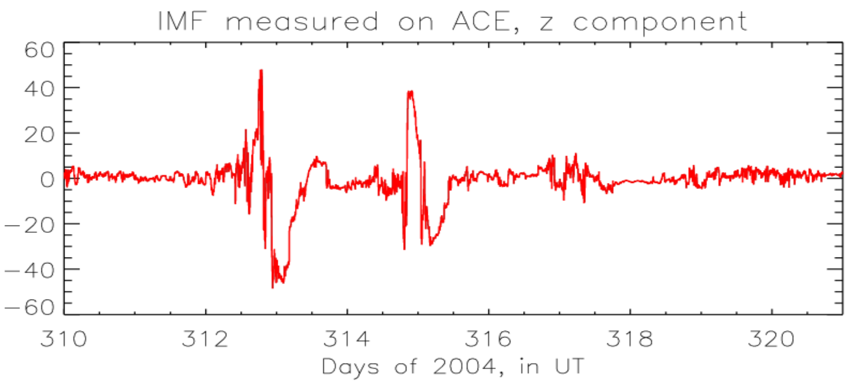
**Definition of an effective scintillation index on the top of a station:**

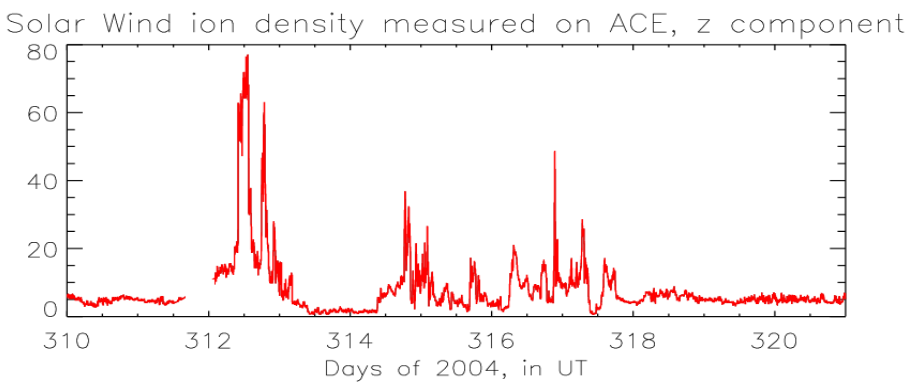
$$\sigma_{\phi}(t) = \sqrt{\sum_{i=1}^{N(t)} \frac{\cos\vartheta(t)}{N(t)} \left(\sigma_{\phi}^{(i)}(t)\right)^2}$$



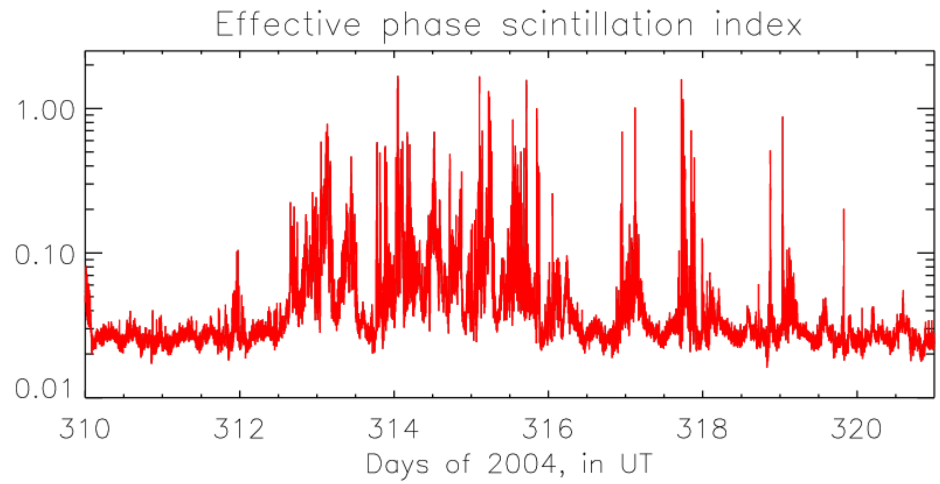
**Solar Wind time series: the z component of the IMF and the ion density, as measured by the satellite ACE (data source: http://spidr.ngdc.noaa.gov/spidr/).**

**Period considered: from November 5 to November 15 of 2004.**



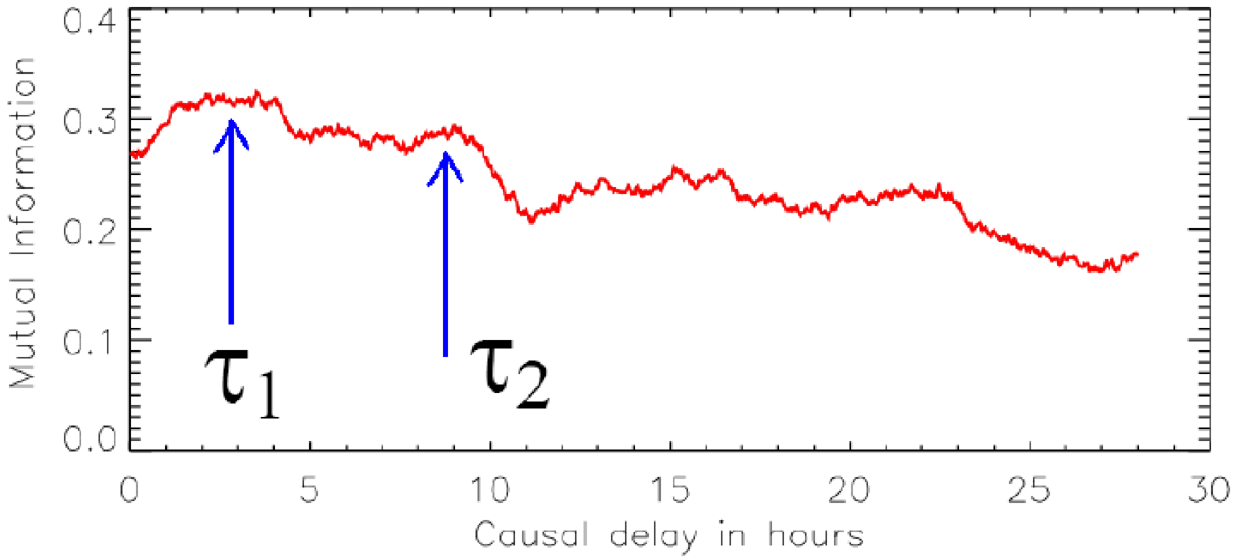




**GPS scintillation data collected with a GSV4004, produced by GPS Silicon Valley, maintained in Tromsø by the University of Bath (UK).**



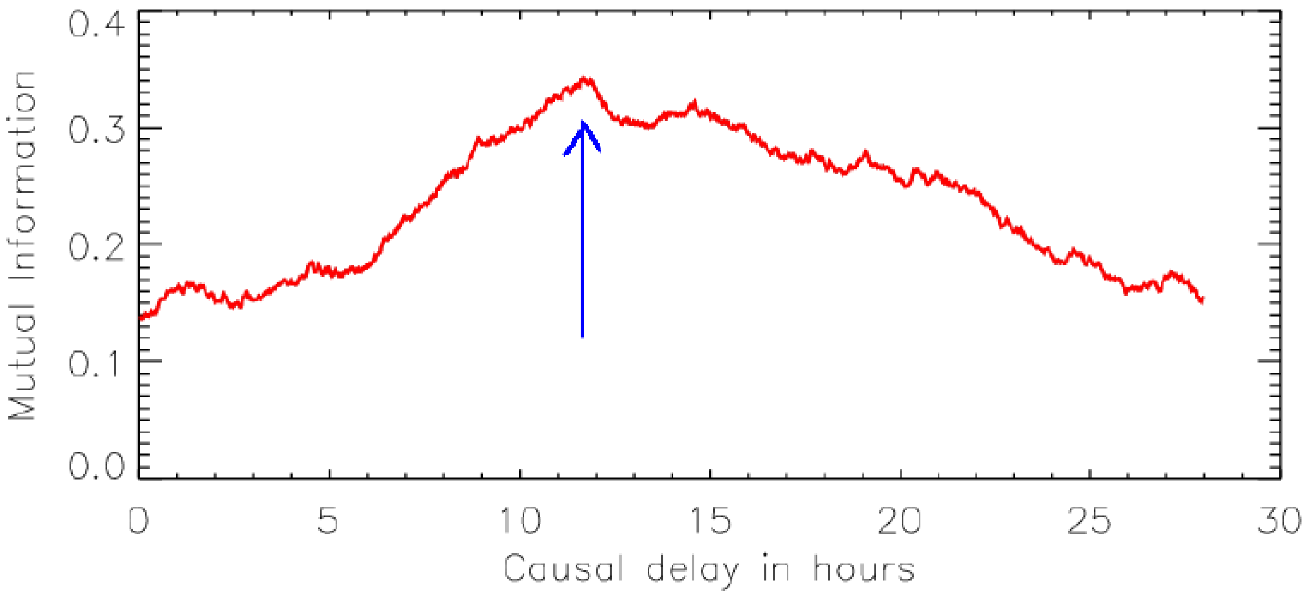
## *$B_z$* delayed mutual information with phase scintillation: the maximum response of sigma-phi to the IMF z-component takes place between 1 and 5 hours, with another peak at 8 hours delay.

$$M\left(B_z\left(t\right),\sigma_\phi\left(t+\tau\right)\right)$$



## $N_i$ delayed mutual information with phase scintillation: the maximum response of sigma-phi to the Solar Wind ion density takes place with a delay of 11-12 hours, while a smaller maximum appears after 15 hours.

$$M\left(N_i\left(t\right), \sigma_\phi\left(t+\tau\right)\right)$$



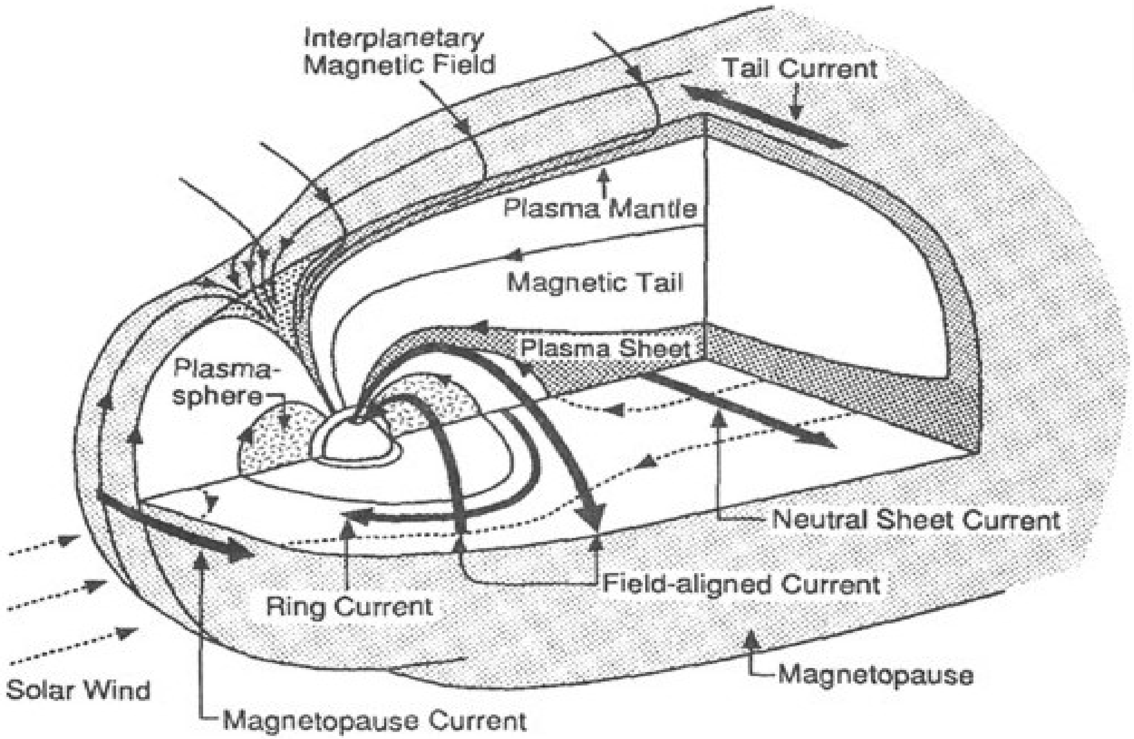
# IT analysis 4/4: storm-substorm relationship via transfer entropy analysis of AL and SYM-H data

An information theory approach to the storm-substorm relationship

P. De Michelis,[1,2] G. Consolini,[3] M. Materassi,[4] and R. Tozzi[1]

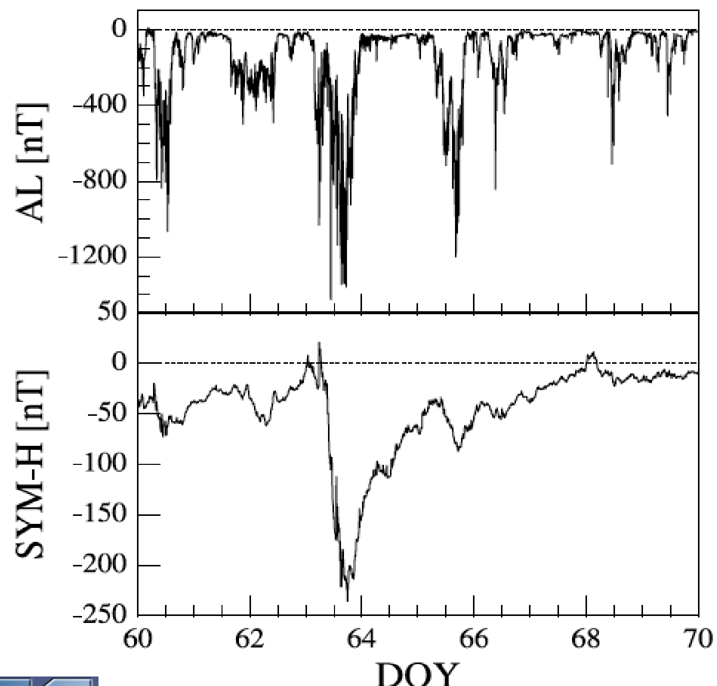
JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 116, A08225, doi:10.1029/2011JA016535, 2011

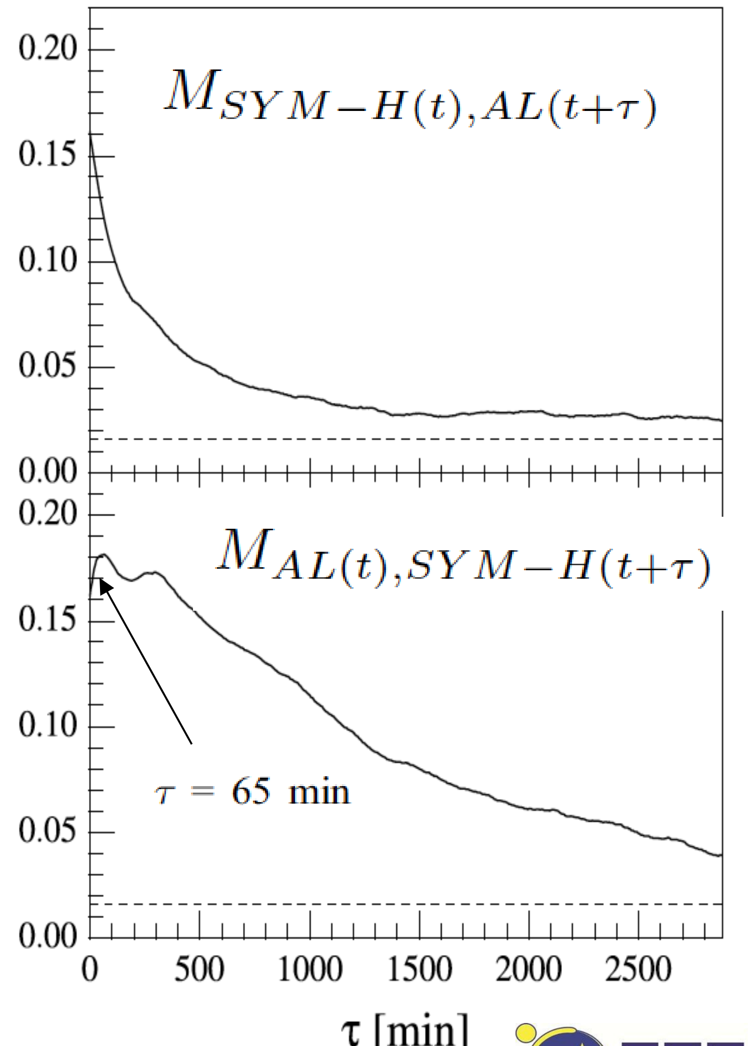
**Akasofu 2004: “The magnetospheric substorm is a fundamental mode of magnetospheric disturbance… When intense substorms occur frequently, a magnetospheric storm develops as their non-linear consequence.”**

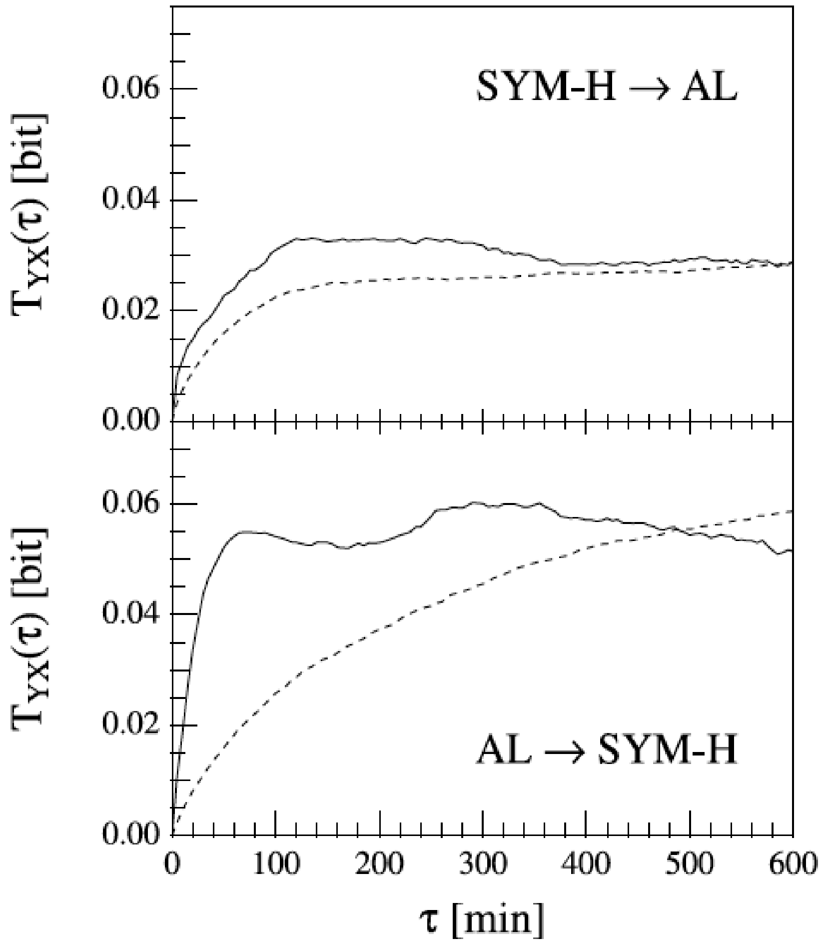



**It has also been found that substorms may occur at times when there is no ring current enhancement or when the ring current is recovering. Thus, substorms appear to be something more than a simple component of a magnetic storm.**

**1981 year of 1' data averaged over 5'. (World Data Center for Geomagnetism (Kyoto, Japan)). Need of "stationary" hypothesis over 365 days. All "averaged"...**

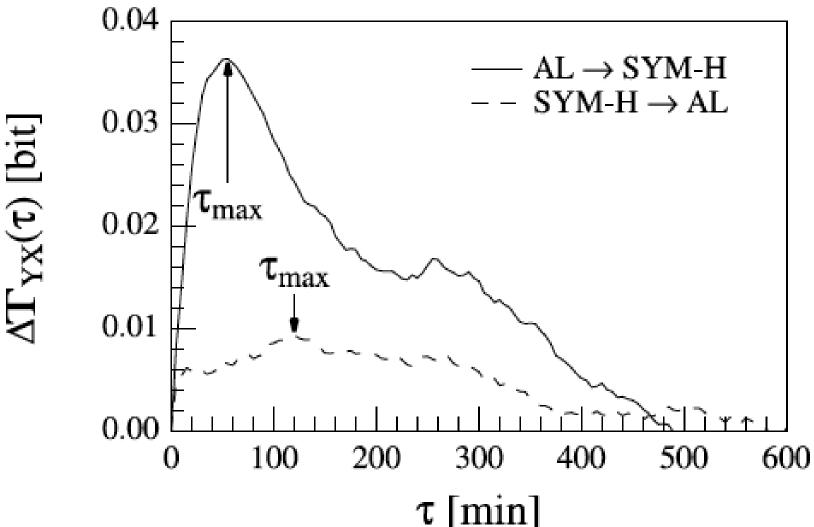




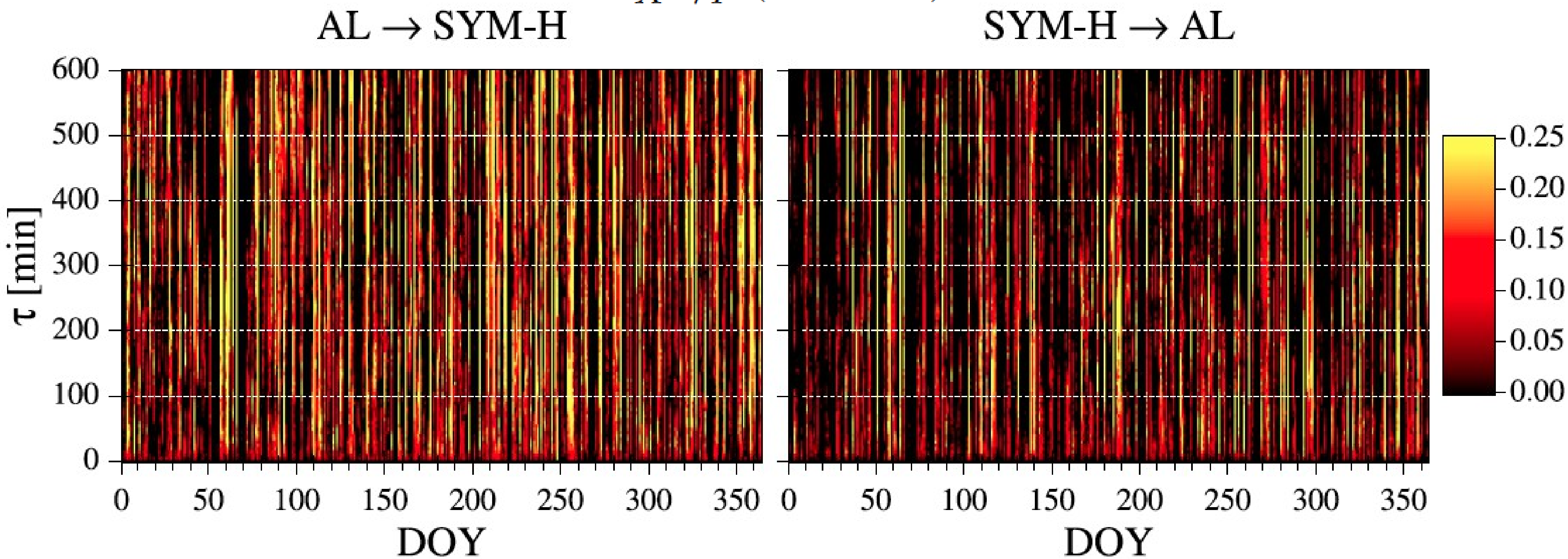

**The reliable part of the transfer entropies assesses that the prevailing influence is from AL to SYM-H, hence a net influence from sub-storms to storms appear to exist**

$$T^{(\mathrm{reli})}_{X\to Y}(t,\tau) = T_{X\to Y}(t,\tau) - T^{(\mathrm{SDT})}_{X\to Y}(t,\tau)$$

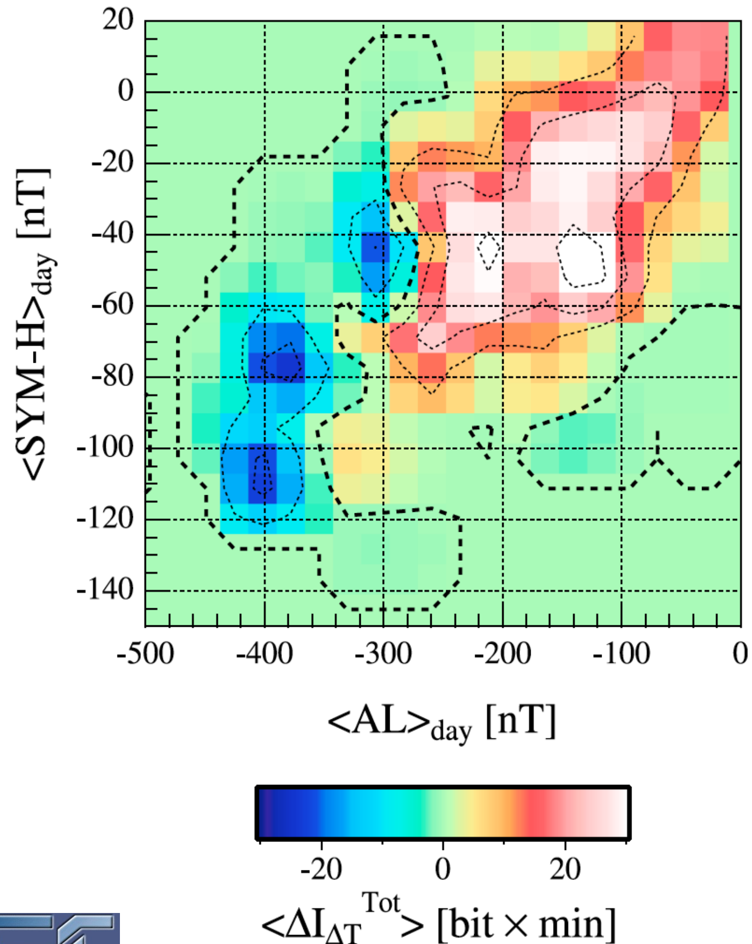


**The integral of the reliable TE over the significative time is 7.82 bit*min for AL2SYM-H and 2.72 bit*min for SYM-H2AL. Meaningful contribution in both directions, which is confirmed by looking at the daily-based TES:**

$$T^{(\mathrm{reli})}_{X \to Y}\,(\mathrm{DOY}, \tau)$$



$$I^{\mathrm{tot}}_{\mathrm{AL}\rightarrow\mathrm{SYM-H}}\left(\mathrm{DOY}\right) = \int_0^{10\mathrm{h}} \left[T^{(\mathrm{reli})}_{\mathrm{AL}\rightarrow\mathrm{SYM-H}}\left(\mathrm{DOY},\tau\right) - T^{(\mathrm{reli})}_{\mathrm{SYM-H}\rightarrow\mathrm{AL}}\left(\mathrm{DOY},\tau\right)\right] d\tau$$

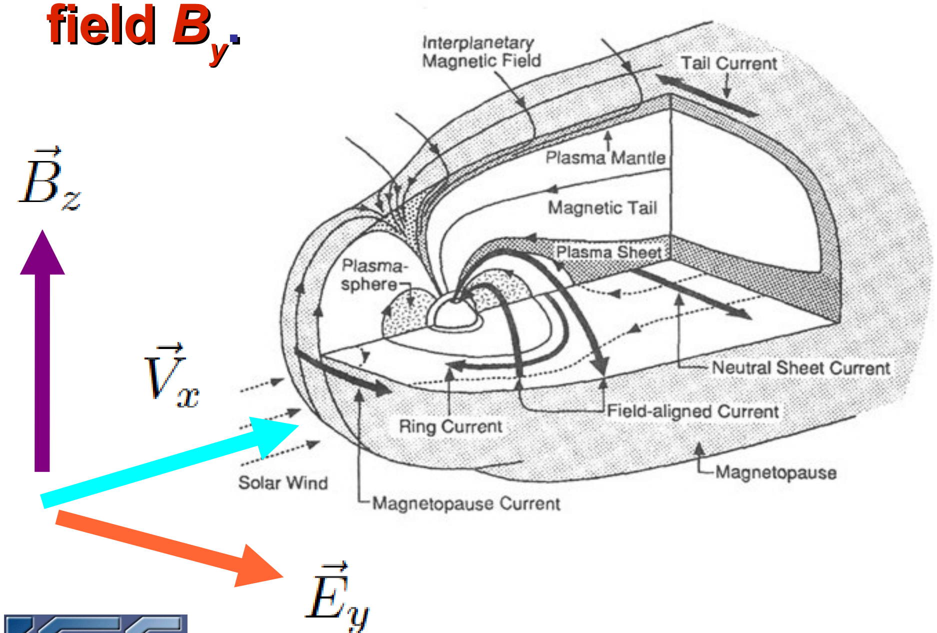


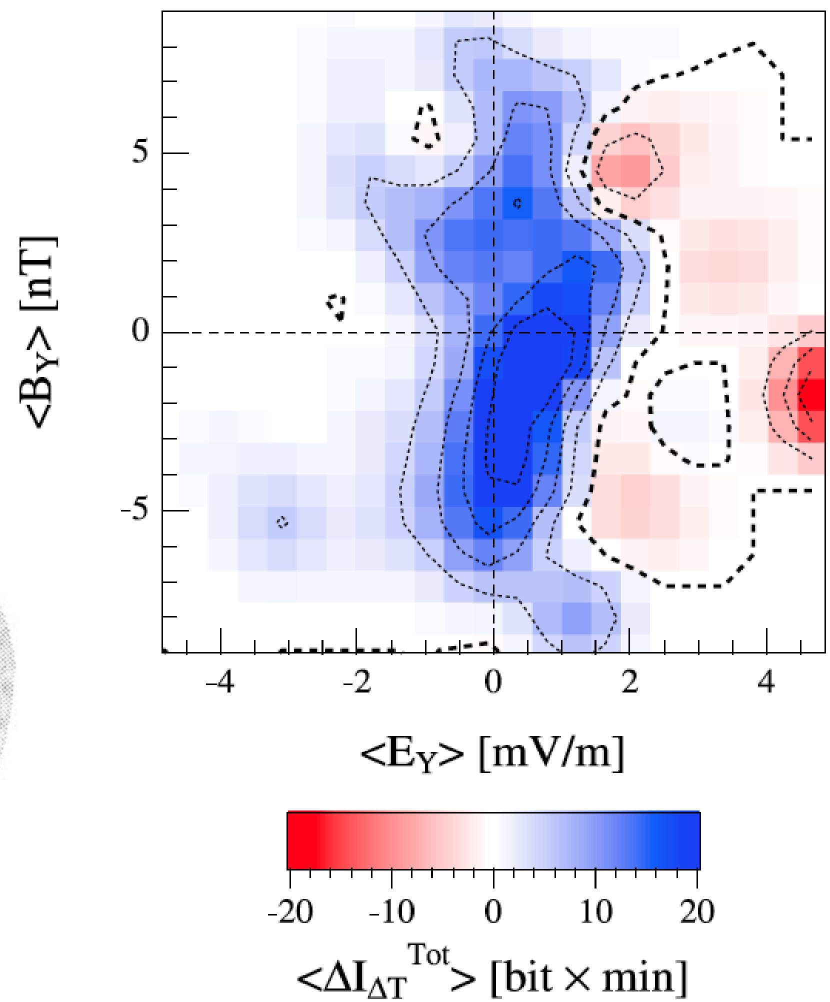

**Total TE exchanged in a day versus day-averaged indices.**

**Searching for evidence of “order parameters” controlling the possible dynamical phase transition.**

**Opposite conditions are not mixed up together, because two distinct dynamical phases appear to exist!**

**Two dynamical phases can be also highlighted with respect to the conditions of the solar wind, described by the daily average of the convection electric field $E_y$ and magnetic field $B_y$.**



The different directionality in the information slave-master relationship between AL and SYM–H may be the signature of different internal and external characters of the observed magnetic storms.

Of course, information theory doesn't give any answer about "why" and "how" this happens in detail, in terms, e.g., of differential equations.

Nevertheless, it provides with important constraints (the "driving direction", the response delay) and important indications (the dynamical phase transition and the domains of the two phases), according to which attempts of modelling might be fruitful.

1. What is information theory;

2. Information theory and coupled processes: mutual information, delayed mutual information, transfer entropy;

3. Information theory vs Space Weather: cascades, near-Earth plasma physics and ionospheric scintillation, storms and substorms;

4. Conclusions and remarks

**Modelling: understanding how to construct dynamical models of interaction between processes starting from the information transfer analysis.**

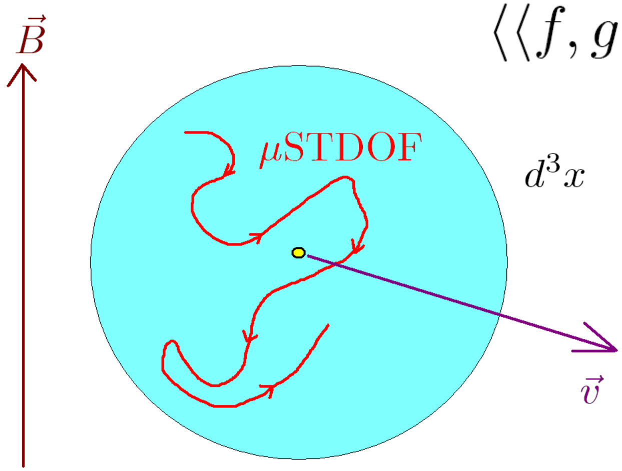
$$\dot{x} = f\left(x, y, \xi\right),$$

$$\dot{y} = g\left(x, y, \eta\right)$$

$$T_{X \to Y}\left(t, \tau\right)$$

**Notice: processes may well be non-Markovian, non-local etc...**

**Entropy-information and dynamics: in complex systems (and the Helio-Geospace is a complex system!) one has the opportunity to conjugate “traditional” dynamics with information theory. The role of Entropy as a quantification of non-precise-knowledge of some degrees of freedom emerges, influencing the dynamics! E.g. the metriplectic framework for dissipative MHD**

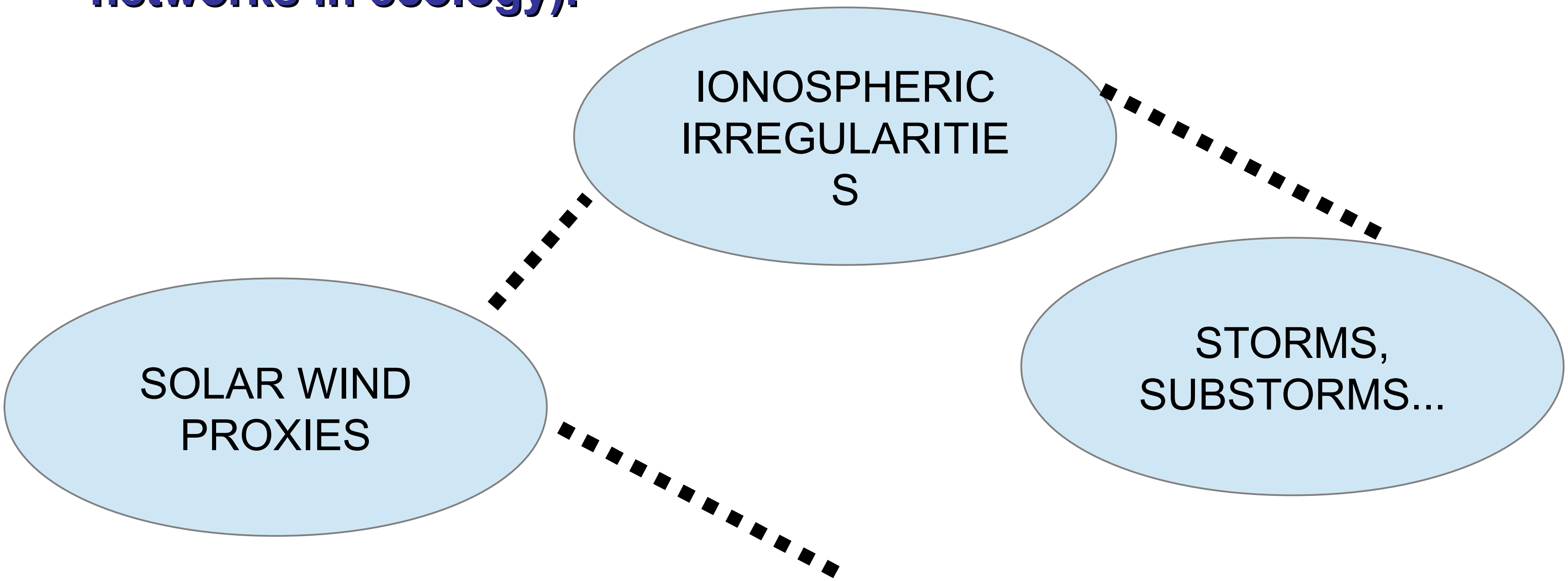



$$\langle\langle f,g\rangle\rangle = \{f,g\} + (f,g) \quad \forall \quad f = f(\psi), \quad g = g(\psi)$$

$$F = H + \alpha S$$

$$\dot{f} = \partial_t f + \langle\langle f,F\rangle\rangle \quad \forall \quad f = f(\psi),$$

**Information networks: complex systems, made out of (many) sub-systems, may be regarded as networks of processes exchanging influence and driving, that will be quantified via information theory quantities (e.g., trophic cascades or networks in ecology).**



Spineto (Italy), June 09, 2012