Analysis of fMRI time-series by entropy measures

Pavol Mikoláš [1], Jan Vyhnánek [1], Antonín Škoch [2], Jiří Horáček [1]

1 Prague Psychiatric Center, 3rd Faculty of Medicine, Charles University, Prague, Czech Republic
2 Department of Diagnostic and Interventional Radiology, Institute for Clinical and Experimental Medicine, Prague, Czech Republic

Correspondence to: Pavol Mikoláš, MD.
Prague Psychiatric Centre,
Ústavní 91, CZ–181 03 Praha 8 - Bohnice, Czech Republic.
e-mail: mikolas@pcp.lf3.cuni.cz

Submitted: 2012-06-15 Accepted: 2012-09-10 Published online: 2012-10-02

Key words: **approximate entropy; block design; brain; computer simulation; event related design; humans; information entropy; magnetic resonance imaging; methods; mutual information; neuroimaging; resting state**

Neuroendocrinol Lett 2012; **33**(5):471–476 **PMID:** 23090262 NEL330512R01

Abstract

Entropy is a measure of information content or complexity. Information-theoretic modeling has been successfully used in various biological data analyses including functional magnetic resonance (fMRI). Several studies have tested and evaluated entropy measures on simulated datasets and real fMRI data. The efficiency of entropy algorithms has been compared to classical methods based on the linear model. Here we explain and summarize entropy algorithms that have been used in fMRI analysis, their advantages over classical methods and their potential use in event-related and block design fMRI.

Abbreviations:

ApEn	- approximate entropy
BOLD	- blood oxygen level-dependent
ER-fMRI	- event-related fMRI
ERP	- event-related potential
fMRI	- functional magnetic resonance imaging
GLM	- general linear model
GRE	- generalized relative entropy
HPEI	- Hilbert phase entropy imaging
HRF	- hemodynamic response function
IT	- inspection time
MI	- mutual information
ROC	- receiver operating characteristic
SNR	- signal-to-noise ratio
TFR	- time frequency representation

INTRODUCTION

fMRI (functional magnetic resonance imaging) is a well established experimental method in modern cognitive neuroscience that has dominated the field over the last two decades. Generally fMRI data have low amplitude and low signal-to-noise ratio (SNR). The need for extraction of reliable functional information has stimulated development of specialized deterministic, statistical and informational algorithms (D'Esposito *et al.* 1999; Fuhrmann Alpert *et al.* 2007). In this review we focus on an alternative approach to fMRI analysis using information theoretic entropy measures.

REVIEW ARTICLE

fMRI AND MECHANISM OF BOLD SIGNAL

The fMRI detects decreases in deoxygenated hemoglobin induced by increased blood flow in areas of increased neuronal activity. The shape of the detected blood oxygen level-dependent (BOLD) signal is represented by the hemodynamic response function (HRF) (Figure 1). HRF is an empirically defined model of the assumed response after a stimulus with a typical shape – a positive BOLD signal detected 4–5 sec after a stimulus. This signal reaches its maximum in about 7 seconds (time-to-peak) and a baseline in approximately 20 seconds (Malonek & Grinvald 1996).

There are three frequently used paradigms for stimulus related fMRI experiments – block design, event-related design (ER-fMRI) and resting fMRI. In the block design, one condition is presented over a discrete period of time (in a 'block') in a repeated or continuous fashion. Resulting time-integrated signal is contrasted with the signal from blocks with different conditions. In ER-fMRI one or more conditions followed by a period of rest are presented in repeated fashion and the detected signal reflects the response to an individual condition rather than a time-integrated response (Donaldson & Buckner 2001). This approach is suitable for infrequent or complex stimuli that cannot be presented in block paradigms. Another kind of experimental design is the no-stimulus mode, or "resting state" (Raichle *et al.* 2001) where subjects are asked to rest with their eyes closed and engage in mind wandering. Data analysis then focuses on detection of organized auto-correlated resting-state networks (Biswal *et al.* 1997; Greicius *et al.* 2003). Analysis of the resting-state has been main focus of recent interest. The range of methods for its analysis has been expanding, and the development of new tools to explore relationships between brain regions is expected.

DEFINITION OF ENTROPY

Entropy quantifies the amount of uncertainty in a system. In other words, entropy is a measure of randomness, or regularity, where regularity is given by low entropy values and randomness by high entropy values. Entropy was first introduced by Shannon (Shannon 1948) in the form of the so-called Shannon entropy (see below). Since then various entropy measures have been developed which quantify the uncertainty of systems by different means.

Entropy is a suitable relative measure for comparing stochastic biological data which are in the form of sampled real-valued signals (i.e. real-valued time series). Unlike moment statistics (mean and variance), entropy does not depend on absolute values of the signal. Instead it reflects the regularity in the distribution of values, or in some cases it measures regularity of consistency within data. Entropy has been successfully applied in biological science for analysis of various physiologic systems such as heart-rate variability, hormone secretion, negative feedback strength, EEG, MRI, and now fMRI (Palus. 1996; Pincus *et al.* 1999; Pincus. 2006; Stam. 2005).

ASSUMPTIONS FOR ENTROPY ANALYSIS OF fMRI

One advantage of entropy measurement over conventional methods is that it requires few assumptions about the nature of hemodynamic responses, underlying neural processes or data itself (de Araujo *et al.* 2003). A common assumption in the classical linear transform model is that the fMRI responses are proportional to local mean neural activity averaged over a period of time (Boynton *et al.* 1996). This infers that the relationship between the stimulus and HRF is linear, in other words more intense stimulation produces stronger response (Worsley & Friston. 1995). This was shown not to be true for short stimulation ER-fMRI paradigms (Vazquez & Noll. 1998). Another common assumption is that the hemodynamic response function has a fixed shape. However, it has been estimated that there is considerable variation in HRF over different regions in individual subjects, across cognitive task paradigms and also between subjects (Miezin *et al.* 2000). The only assumptions for BOLD analysis using entropy are made concerning the number of system states (eg. activation and rest) and the statistical independence of time-series (Sturzbecher *et al.* 2009).

ALGORITHMS OF ENTROPY ANALYSIS IN fMRI

In order to apply some entropy measures upon a real-valued series, N discrete intervals (or amplitude levels) $I_{k=1..N}$ are selected so that every value from a series belongs to some interval I_j (Figure 2). We define the probability p_k that a value from a series X belongs to k-th interval I_k:

$$p_k = \frac{No.\ values\ \ of\ \ series\ \ within\ \ I_k}{No.\ all\ \ values\ \ of\ \ series}$$

Shannon entropy quantifies the randomness with which the values are distributed into intervals I_k.

Shannon entropy H of a series X of length N is defined as:

$$H(X) = -\sum_{k=1}^{N} p_k \cdot log\,(p_k)$$

Intuitively, Shannon Entropy is high if values of a series X cover intervals I_k rather uniformely and low if values from a series X belong to only a few intervals I_k. E.g. if all values from a series belong to an interval I_k, then $p_k=1$ and $H(X)=0$ which is the lowest possible value of $H(X)$. This fits with the fact that the series is not random at all.

De Araujo *et al.* (2003) were first to use Shannon entropy to analyze fMRI time-series. Visual flashing light and bilateral motor (finger tapping) stimuli were presented to 9 healthy volunteers in block and ER manner. The results of block paradigm were analyzed by cross-correlation coefficient mapping with a boxcar reference function. ER fMRI results were analyzed by two independent methods: cross-correlation between each voxel's time-course and a lagged gamma function; and Shannon entropy dependent on time. Entropy was calculated over two different time-windows, reflecting activation and rest. Statistical maps were obtained by correlating entropy values with a simulated sawtooth function reflecting alternating stimulus and rest time-windows. Window parameters were optimized for analysis of visual and motor paradigms. Shannon entropy was shown to be an effective method of ER-fMRI analysis, having several advantages over classical cross-correlation, such as better consistency with decreasing signal-to-noise ratio and model independency.

This approach was extended by Sturzbecher *et al.* using **Tsallis entropy** (Tsallis. 1988). For a series X of discrete values the Tsallis entropy of order q is defined as:

$$H_q(X) = \frac{1}{q-1} \cdot \left(1 - \sum_{k=1}^{N} p_k^q \right)$$

where q is a parameter which influences the characteristics of the measure. For $q \rightarrow 1$ Tsallis entropy tends to follow the pattern of Shannon entropy. For other values the interpretation is not clear. As in the case of Shannon entropy, in this measure low values of entropy correspond to series with values distributed among fewer intervals while high values of entropy correspond to series with values distributed randomly among high number of intervals.

Tsallis entropy was calculated for two different time-windows (activation and rest). The method was tested on simulated data and real ER-fMRI data (visual and motor paradigm) using several combinations of input parameters and compared to general linear model using receiver operating characteristic (ROC) curves. For simulated HRFs Tsallis entropy was more stable with both changing signal-to-noise ratio and HRF delays than the general linear model (GLM). It was also more sensitive in detecting activation than Shannon entropy (Sturzbecher *et al.* 2009; Tedeschi *et al.* 2004).

A similar approach was introduced by Cabella *et al.* (2009). Authors analyzed ER-fMRI simulated and real data using generalized Kullback-Leiber distance D_q, also referred to as the **generalized relative entropy (GRE)**. ER BOLD signal time series were divided into two time-windows W_1 and W_2 reflecting periods of signal and rest respectively. D_q represents the distance between probability functions for the two states. Calculation of D_q requires two input parameters: L is the number of amplitude levels and q is the Tsallis q parameter. The choice of suitable input parameters was evaluated on simulated data with variable SNR by ROC curves. Real data from a finger-tapping paradigm were analyzed and statistical maps were constructed with different cutoff values. Authors concluded that D_q is a suitable method for fMRI analysis, although they have not compared the method to any conventional methods to investigate possible specific advantages or disadvantages in robustness.

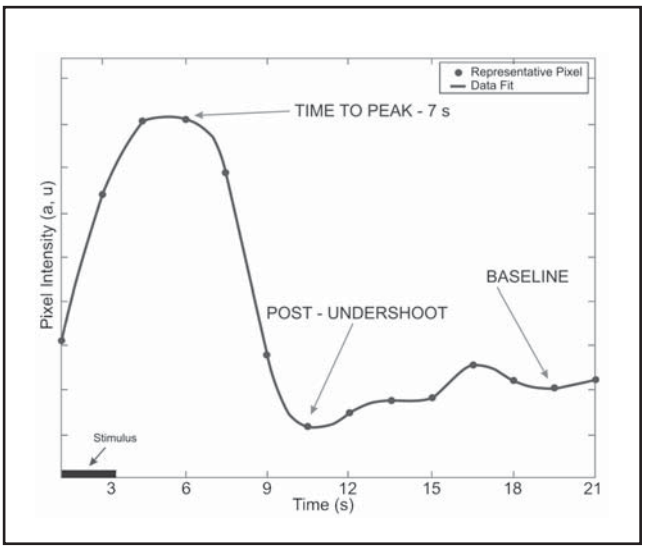

Fig. 1. Assumed shape of the haemodynamic response function (HRF). It reaches peak at about 7s after the stimulus onset (time-to-peak) and returns to baseline at about 20 sec. *Note.* Adapted from "Shannon entropy applied to the analysis of event-related fMRI time series" by de Araujo *et al.* (2003).

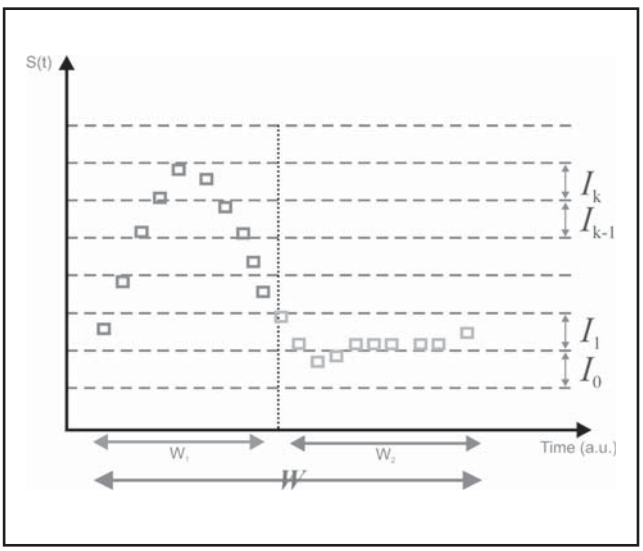

Fig. 2. ER-fMRI analysis by Shannon entropy. Each individual voxel's time series was divided into periods of stimulation (W_1) and rest (W_2). The amplitude of signal was divided into k intervals or possible system states (I_0 - I_k) so that every amplitude level belongs to certain interval I_k. Entropy was calculated for each time-window separately. *Note.* Adapted from "Shannon entropy applied to the analysis of event-related fMRI time series" by de Araujo *et al.* (2003).

Andino *et al.* (2000) developed the **Renyi number** – a combination of Renyi entropy and time frequency representation (TFR) of a signal in order to differentiate between signal and noise in ER-fMRI, intracranial event-related potentials (ERPs) and EEG time-series. TFR divides the signal waveform into specific components (or 'energy spots') according to the power spectral density. Activation is represented by a higher organization of the signal, and therefore fewer components, whereas noise is represented by multiplicity of components. Renyi entropy is then applied over the basis of TFRs. In a sequential motor fMRI paradigm, this approach was shown to more effectively distinguish between activation and noise when compared to correlation analysis. Moreover, RE was effective in detecting highly organized signals in the motor cortex during a simple index finger response task measured by ERPs in 2 epileptic patients as well as in the detection of activation in a simple visual task measured by ELECTRA, an EEG inverse solution that estimates three dimensional potential inside the brain.

Adaptive entropy rates quantify how precisely the k-th value can be predicted from past information. This approach was used by Fisher *et al.* (2001) to analyze motor, auditory and visual paradigms and compared to the general linear model (Worsley & Friston. 1995) and mutual information (Tsai *et al.* 1999). The study examined whether the values within the areas of interest could be predicted better in the case of using the information of (a) k–1 previous values and (b) k–1 previous values combined with the additional information on experiment protocol. Both predictions were made by linear combinations of preceding signal values, resp. signal values and protocol values. Adaptive entropy rates were found to have minor advantages over GLM and mutual information (MI, see below) but authors suggest its potential where it is difficult to model signals *á priori*.

In order to utilize phase information from fMRI time-series, Liao *et al.* (2010) proposed a new method called **Hilbert phase entropy imaging (HPEI)**. It uses Hilbert phase transformation to determine phase differences between task-state and control-state. If the voxel's states are synchronized, the distribution of phase differences is a peaked distribution (Laird *et al.* 2002). Shannon entropy was used to distinguish between a peaked distribution with low entropy value and rather uniform distribution with high entropy value.

Tab. 1. Summary of information theoretic entropy measures and mutual information in fMRI analysis.

Approach	Principle	Application	Comparison with standard methods	References
Shannon entropy	Quantifies the regularity with which values are distributed into intervals	ER-fMRI visual and motor paradigm	Better consistency with decreasing SNR than cross-correlation	(Cabella *et al.* 2009)
Tsallis entropy	Quantifies the regularity with which values are distributed into intervals, Shannon entropy modified by the q-parameter	ER-fMRI visual and motor paradigm, simulated data	More stable with changing SNR ratio than GLM, more sensitive in detecting activation than Shannon entropy	(Sturzbecher *et al.* 2009)
Adaptive entropy rates	Quantifies how precisely a value can be predicted from past information	ER-fMRI motor, auditory and visual paradigm	Minor advantages over GLM and mutual information	(Tsai *et al.* 1999), (Fisher *et al.* 2001)
General relative entropy	Calculates distance between probability functions in two time-windows	ER-fMRI motor paradigm	no comparison	(Cabella *et al.* 2009)
Hilbert phase entropy	Uses Hilbert phase transform to determine phase differences between task-state and control-state	ER-fMRI and block-design visual and motor paradigm, simulated data	More effective than SPM and Laird's method (Laird *et al.* 2002)	(Liao *et al.* 2010)
Renyi number	Entropy measure applied over time frequency representation of a signal	Sequential motor paradigm	Same effectiveness as correlation. May be applied to experiments where on/off conditions are not available	(Andino *et al.* 2000)
Approximate entropy	Quantifies the regularity of patterns contained in a time-series	IT visual information processing task	Novel approach to evaluate decrease in signal complexity associated with lifelong cognitive change	(Sokunbi *et al.* 2011)
Mutual information	Quantifies the mutual dependency of two time series	Simulated data and real block-design motor paradigm	More stable threshold than cross-correlation	(Tsai *et al.* 1999)

ER-fMRI - event related functional MRI; SNR - signal-to-noise ratio; GLM - general linear model; IT - inspection time

The efficiency of HPEI was evaluated on simulated data with variable parameters of delay, signal-to-noise ratio and shape of HRF; and on real experiments using visual and motor paradigms. In all conditions HPEI was shown to be more effective than SPM and Laird's method (Laird *et al.* 2002; Liao *et al.* 2010).

A novel approach to evaluate entropy changes in temporal fMRI signal was used by Sokunbi *et al.* (2011). Authors investigated entropy changes associated with lifelong cognitive change. fMRI signal obtained from 40 subjects during the inspection time (IT) visual information processing task was analyzed by **approximate entropy (ApEn)**. High ApEn values in regions especially involved in visual processing were associated with better cognitive performance.

Unlike other entropy measures used in fMRI analysis (see above), ApEn is a suitable method for temporal fMRI analysis as it does not require differentiating time-series into time-windows. It is defined as:

$$ApEn\,(m,r) = \varphi^{m}(r) - \varphi^{m+1}(r)$$

$\varphi^{m}(r)$ quantifies the similarity of sequences of length *m* and is defined as:

$$\varphi^{m}(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} log\,(C_i^{m})$$

where *N* is the length of the time-series and C_i^m is the number of patterns of length *m* which are similar to the pattern beginning at the position *i*.

ApEn quantifies the regularity of the patterns contained in a time-series. More specifically it measures the probability that two randomly chosen patterns of length *m+1* will be similar given that they were similar on their first *m* time points. In other words it measures the stability with which two similar patterns rest similar after extending them of one time point. Intuitively, if ApEn of the time-series is low, the similarity of patterns is stable in time and therefore we can consider the series to be more regular than a series with high ApEn (Pincus *et al.* 1991).

Mutual information (MI) is an information theoretic approach that calculates the predictability of one mutually dependent time series given the information from using entropy. In ER-fMRI, MI is calculated between fMRI temporal responses and the experimental. Therefore, unlike entropy measures, MI requires assumptions about the experimental protocol. This approach was introduced by Tsai *et al.* (1999) and further optimized by Tedeschi *et al.* (2005) in the form of GMI (generalized mutual information). GMI was shown to be effective in differentiating between between activation and rest in simulated and real fMRI data (motor paradigm). GMI was shown to be less prone to thresholding than cross-correlation. Use of MI was also extended to resting state fMRI functional connectivity analysis in a range of algorithms. For example by calculating MI between the individual voxel time-series (Benjaminsson *et al.* 2010) or the averaged sub-region time-series (Lizier *et al.* 2011) or in sub-regions frequency domain (Salvador *et al.* 2007) this approach has advantage of detecting non-linearities and could be extended to account for directional relationships (Lizier *et al.* 2011). However, in general it was found to have minor benefits when compared to linear correlation in detecting resting-state functional connectivity (Hlinka *et al.* 2011). MI was not widely adopted for ER-fMRI analysis either.

CONCLUSION

Entropy measures have been consistently shown to be suitable methods for evaluation of ER and temporal fMRI (Table 1). In comparison with standard methods they offer model independence and to some extent better outcomes with changing signal-to-noise ratio. However, these methods have not been widely adopted for ER-fMRI experiments, which may partially be due to high demands on computation power. On the other hand, suitability of entropy measures for emerging fMRI experimental paradigms including resting-state fMRI has not been fully investigated.

ACKNOWLEDGEMENTS

This work was supported by grants GACR P304/12/G069, IGA MHCR NT/13843, "2009 J.Horacek NARSAD Independent Investigator Award" and research project MZ0IKEM2005.

REFERENCES

1 Andino SLG, Menendez RGD, Thut G, Spinelli L, Blanke O, Michel CM et al (2000). Measuring the complexity of time series: An application to neurophysiological signals. Human Brain Mapping **11**: 46–57.
2 Benjaminsson S, Fransson P, Lansner A (2010). A novel model-free data analysis technique based on clustering in a mutual information space: application to resting-state FMRI. Front Syst Neurosci **4**: 34.
3 Biswal BB, VanKylen J, Hyde JS (1997). Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. Nmr in Biomedicine **10**: 165–170.
4 Boynton GM, Engel SA, Glover GH, Heeger DJ (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. Journal of Neuroscience **16**: 4207–4221.
5 Cabella BCT, Sturzbecher MJ, de Araujo DB, Neves UPC (2009). Generalized relative entropy in functional magnetic resonance imaging. Physica A-Statistical Mechanics and Its Applications **388**: 41–50.
6 D'Esposito M, Zarahn E, Aguirre GK (1999). Event-related functional MRI: Implications for cognitive psychology. Psychological Bulletin **125**: 155–164.
7 de Araujo DB, Tedeschi W, Santos AC, Elias J, Neves UPC, Baffa O (2003). Shannon entropy applied to the analysis of event-related fMRI time series. Neuroimage **20**: 311–317.
8 Donaldson DI, Buckner RL (2001) Effective paradigm design. In: Jezzard P, Matthews PM, Smith SM editors. Functional MRI: An introdction to methods. New York: Oxford University Press. p. 177–195.
9 Fisher JW, Cosman ER, Wible C, Wells WM (2001) Adaptive Entropy Rates for fMRI Time-Series Analysis. Med Image Comput Comput Assist Interv. MICCAI 2001, pp. 905–912.

10 Fuhrmann Alpert G, Sun FT, Handwerker D, D'Esposito M, Knight RT (2007). Spatio-temporal information analysis of event-related BOLD responses. Neuroimage **34**: 1545–1561.
11 Greicius MD, Krasnow B, Reiss AL, Menon V (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. Proceedings of the National Academy of Sciences of the United States of America **100**: 253–258.
12 Hlinka J, Palus M, Vejmelka M, Mantini D, Corbetta M (2011). Functional connectivity in resting-state fMRI: Is linear correlation sufficient? Neuroimage **54**: 2218–2225.
13 Laird AR, Rogers BP, Carew JD, Arfanakis K, Moritz CH, Meyerand ME (2002). Characterizing instantaneous phase relationships in whole-brain fMRI activation data. Human Brain Mapping **16**: 71–80.
14 Liao W, Chen HF, Pan ZY (2010). Hilbert phase entropy imaging of fMRI time series. Computer Methods in Biomechanics and Biomedical Engineering **13**: 121–133.
15 Lizier JT, Heinzle J, Horstmann A, Haynes JD, Prokopenko M (2011). Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. Journal of Computational Neuroscience **30**: 85–107.
16 Malonek D& Grinvald A (1996). Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: Implications for functional brain mapping. Science **272**: 551–554.
17 Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage **11**: 735–759.
18 Palus M (1996). Coarse-grained entropy rates for characterization of complex time series. Physica D **93**: 64–77.
19 Pincus SM (2006). Approximate entropy as a measure of irregularity for psychiatric serial metrics. Bipolar Disorders 8: 430–440.
20 Pincus SM, Gladstone IM, Ehrenkranz RA (1991). A Regularity Statistic for Medical Data-Analysis. Journal of Clinical Monitoring **7**: 335–345.
21 Pincus SM, Hartman ML, Roelfsema F, Thorner MO, Veldhuis JD (1999). Hormone pulsatility discrimination via coarse and short time sampling. American Journal of Physiology-Endocrinology and Metabolism **277**: E948–E957.
22 Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001). A default mode of brain function. Proceedings of the National Academy of Sciences of the United States of America **98**: 676–682.
23 Salvador R, Martinez A, Pomarol-Clotet E, Sarro S, Suckling J, Bullmore E (2007). Frequency based mutual information measures between clusters of brain regions in functional magnetic resonance imaging. Neuroimage **35**: 83–88.
24 Shannon CE (1948). A Mathematical Theory of Communication. Bell System Technical Journal **27**: 379–423.
25 Sokunbi MO, Staff RT, Waiter GD, Ahearn TS, Fox HC, Deary IJ et al (2011). Inter-individual Differences in fMRI Entropy Measurements in Old Age. Ieee Transactions on Biomedical Engineering **58**: 3206–3214.
26 Stam CJ (2005). Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. Clinical Neurophysiology **116**: 2266–2301.
27 Sturzbecher MJ, Tedeschi W, Cabella BCT, Baffa O, Neves UPC, de Araujo DB (2009). Non-extensive entropy and the extraction of BOLD spatial information in event-related functional MRI. Physics in Medicine and Biology **54**: 161–174.
28 Tedeschi W, Muller HP, de Araujo DB, Santos AC, Neves UPC, Erne SN et al (2004). Generalized mutual information fMRI analysis: a study of the Tsallis q parameter. Physica A-Statistical Mechanics and Its Applications **344**: 705–711.
29 Tedeschi W, Muller HP, de Araujo DB, Santos AC, Neves UPC, Erne SN et al (2005). Generalized mutual information tests applied to fMRI analysis. Physica A-Statistical Mechanics and Its Applications **352**: 629–644.
30 Tsai A, Fisher JW, Wible C, Wells WM, Kim J, Willsky AS (1999). Analysis of functional MRI data using mutual information. Medical Image Computing and Computer-Assisted Intervention, Miccai'99, Proceedings **1679**: 473–480.
31 Tsallis C (1988). Possible Generalization of Boltzmann-Gibbs Statistics. Journal of Statistical Physics **52**: 479–487.
32 Vazquez AL& Noll DC (1998). Nonlinear aspects of the BOLD response in functional MRI. Neuroimage **7**: 108–118.
33 Worsley KJ& Friston KJ (1995). Analysis of Fmri Time-Series Revisited - Again. Neuroimage **2**: 173–181.