Permutation Entropy Analysis of Dynamical Turbulence in an MHD Plasma Wind Tunnel and the Solar Wind

P.J. Weck, D.A. Schaffner, and M.R. Brown
Swarthmore College, Swarthmore, PA 19081

R.T. Wicks
NASA Goddard Space Flight Center, Greenbelt, MD 20771

The permutation entropy of Bandt and Pompe and the Jensen-Shannon statistical complexity are used to analyze turbulence in the MHD plasma wind tunnel of the Swarthmore Spheromak Experiment. MHD wind-tunnel turbulence is compared to turbulence data from the *Wind* spacecraft and the Large Plasma Device using these metrics, and it is found that the wind-tunnel configuration exhibits more truly turbulent dynamics than LAPD but less turbulent dynamics than those in the solar wind. This paper presents the first use of this statistical analysis tool on solar wind plasma, as well as on an MHD turbulent experimental plasma.

I. Introduction

Since Bandt and Pompe introduced their probability distribution based on ordinal patterns in arbitrary time series in 2002 [3], their methodology has found a wide variety of applications, from tracking the effects of anesthetic drugs on the brain [10, 13, 17] to informing economic policy [4, 27, 28], among many others [11, 20–23]. In 2007, Rosso *et al* applied the ordinal pattern distribution of Bandt and Pompe to a time series analysis using the complexity-entropy plane, or CH plane, capable of differentiating between periodic, chaotic, and stochastic systems [19]. The CH plane has been used to determine the statistical character of fluctuations in several physical systems, including magnetic flux ropes [8] and electron heat transport [16] in magnetized plasmas. However, to our knowledge the approach has yet to be extended to the study of dynamical plasma turbulence, either in the solar wind or in a purely dynamical laboratory plasma. The purpose of the present paper is to fill this gap in the understanding of the CH character of plasma systems, and provide one further point of comparison between the MHD wind tunnel configuration and the solar wind.

We compare the CH positions of MHD wind tunnel, solar wind, and drift-wave plasmas, as well as model chaotic systems and stochastic processes, and find that although more stochastic than the edge plasma in the Large Plasma Device and comparable in entropy and complexity to known stochastic processes, MHD wind tunnel data is both less entropic and more complex than the solar wind. This suggests that further steps must be taken before the wind tunnel configuration can generate the fully developed turbulence seen in the solar wind data. Both larger confinement volumes and longer lifetimes are potential avenues for improving the correspondence between solar wind and MHD wind tunnel positions on the CH plane.

The MHD wind tunnel configuration of the Swarthmore Spheromak Experiment (SSX) consists of a plasma gun which fires a spheromak of magnetized plasma into an ~ 1 meter long cylindrical copper flux conserver, as described in [25] and [6]. Probes embedded in the chamber collect data on turbulent fluctuations in $\dot{B}$ as the plasma evolves down the length of the tube, eventually relaxing into a Taylor state [9]. After injection the plasma is completely dynamical, as there is no guide or vacuum field in the body of the chamber. By varying the amount of magnetic flux through the core of the gun, or "stuffing flux", between 0.0 and 1.5 mWb, the magnetic helicity of the injected plasma can be finely controlled, as described in [6]. Magnetic helicity corresponds to the degree of twistedness in the magnetic field, so varying injected helicity affects the resulting turbulent dynamics of the plasma as it evolves towards a relaxed Taylor state. In our CH plane analysis of turbulence in the SSX plasma source, $\dot{B}$ time series for a variety of injected helicities are used.

Multi-day magnetic fluctuation records from both fast and slow solar wind streams measured by the *Wind* spacecraft were selected for comparison. Fast streams are characterized by high velocities, low densities, and few large-scale structures while slow streams are slower, denser, and more slightly more structured. *Wind* is positioned near 1 AU, and is equipped to collect high cadence (3 s) magnetic field vector data. The magnetic vector component antiparallel to the direction of mean field (away from the sun) are used in our analyzes. Magnetic fluctuations are on the order of 10 nT in both the fast and slow stream signals.

II. Permutation Entropy and the CH Plane

The permutation entropy and thereby the CH plane position of an arbitrary time series is defined in terms of a window length called the embedding dimension n. The embedding dimension determines the size of patterns investigated in calculating the entropy and complexity of the series. The instances of each ordinal patterns of that size are counted in order to associate an ordinal pattern probability distribution with the time series, from which the calculation of entropy and complexity is straightfor-

ward.

For embedding dimension n, the probability distribution introduced by Bandt and Pompe consists of the frequencies of occurrence of all possible length n ordinal patterns in segments of n consecutive terms from an arbitrary time series [3]. In their methodology, a length n ordinal pattern is defined for a segment $s = (x_t, x_{t+1}, \ldots, x_{t+(n-1)})$ of the time series as the permutation π of the index set $\{0, 1, \ldots, n-1\}$ corresponding to the ranking of the x_i in ascending order, namely $x_{\pi_t} < x_{\pi_{t+1}} < \ldots < x_{\pi_{t+(n-1)}}$. In order to guarantee a unique result, if $x_i = x_j$ where $i < j$, then in the ranking $x_i < x_j$. For example, if $x_0 = 5$, $x_1 = -2$, and $x_2 = 0.33$ are three consecutive terms in the time series, then since $x_1 < x_2 < x_0$, the ordinal pattern for this segment is the permutation $\pi = (1, 2, 0)$. Given a time series of length L, the corresponding ordinal pattern probability distribution $P = \{p(\pi)\}$ is defined in terms of all $L - n + 1$ length n segments s in the series and all $n!$ permutations π of order n by

$$p(\pi) = \frac{|\{s : s \text{ has ordinal pattern } \pi\}|}{L - n + 1}. \tag{1}$$

where $|\ldots|$ denotes cardinality. The permutation entropy PE is defined as Shannon's information entropy for this ordinal pattern probability distribution, or

$$PE = -\sum^{n!} p(\pi) \log p(\pi) \tag{2}$$

where the log is base two.

Instead of considering consecutive points in calculating the ordinal pattern probability distribution for a time series, an embedding delay τ can be used to sample ordinal patterns on a larger time scale, thereby placing a lower limit on the size of structures investigated in the frequency spectra and focusing the analysis on a particular temporal scale. Embedding delays can be implemented as a simple sub-sampling of data in which only L/τ values of the time series are considered, as in [16] and [8], or all τ of these portions of the original time series can be used as in [2], henceforth referred to as the length-preserving method. For example, for an embedding delay $\tau = 10$ using the first method, a new time series X' of length $L' = \frac{1}{10}L$ is generated by selecting every tenth value of the original series X and the ordinal pattern probability distribution calculated for that series in the usual manner. In the length-preserving method, segments $(x_t, x_{t+10}, \ldots, x_{t+10(n-1)})$ of X are used to calculate the ordinal pattern probability distribution, where t runs from 1 to $L-10(n-1)$, thereby including the 9/10ths of the dataset thrown out in the first method. Which method is used depends in part on the length of the record in question. Unless $L' \gg n!$, the first method may not yield reliable statistics [8], and the length-preserving method thus appears preferable.

While the permutation entropy quantifies the randomness in an arbitrary time series, a measure of statistical complexity such as the Jensen-Shannon complexity is required to quantify the degree of correlated structure in the time series [19]. The Jensen-Shannon complexity, or C_{JS}, is a functional of the discrete distribution P of N probabilities associated with the time series. Once normalized such that $0 \leq C_{JS} \leq 1$,

$$C_{js}[P] = -2 \frac{S\left[\frac{P+P_e}{2}\right] - \frac{1}{2}S[P] - \frac{1}{2}S[P_e]}{\frac{N+1}{N}\log(N+1) - 2\log(2N) + \log(N)} H[P] \tag{3}$$

Where S is the Shannon entropy, H is the normalized Shannon entropy $S/S_{\max}$, and $P_e = \{\frac{1}{N}, \ldots, \frac{1}{N}\}$ is the uniform distribution. When analyzing time series using the CH plane methodology of Rosso *et al*, this measure of statistical complexity is evaluated by associating with the time series the ordinal pattern probability distribution of Bandt and Pompe, so that $N = n!$, $S[P] = PE$ and $H[P] = PE_{\text{norm}} = PE/\log(n!)$. The statistical nature of time series is then evaluated by comparing their positions on the CH plane, $PE_{\text{norm}} \times C_{JS}$.

In interpreting the CH position of experimental data, it is often useful to compare them to model stochastic processes and chaotic systems. To illustrate the regions of the plane corresponding to paradigmatically stochastic and chaotic dynamics, stochastic fractional Brownian motion (fBm) and chaotic Henon, skew tent, and logistic maps are included in the $n = 5$ CH plane shown in Figure 1. The range of fBm points displayed was generated by varying the corresponding Hurst exponent between 0 and 1, thereby scanning the degree of correlation between succsessive increments of motion from strong negative correlations to positive correlations. A time series generated by a sine function is included as well, illustrating the low entropy, low complexity domain occupied by simple mathematical functions. Note that pure white noise occupies the $PE_{\text{norm}} = 1$, $C_{JS} = 0$ corner of the plane. Also shown in Figure 1 are curves demarcating the minimum and maximum complexity bounds of the CH plane. The nature of the dependence of C_{JS} on H constrains the possible values of the former as a function of the latter to fall between these curves, as described in [14] and [5].

III. CH comparison of SSX and WIND data

The $\dot{B}$ fluctuation signals used in our CH plane analysis of SSX data were recorded by a 16-channel, 3-direction, single-loop pickup coil probe embedded in the midplane of the cylindrical wind tunnel, with a 65 MHz sampling rate and 14 bit dynamic range. 20 μs windows were used in our analysis, corresponding to 1,300-value time series. Approximately 40 shots were taken at each of 8 stuffing fluxes between 0.0 and 1.5 mWb, corresponding to injected helicities ranging from 0 to 7×10^{-5} Wb^2 [6]. In this manner, over 15,000 magnetic fluctuation time series were generated, over a range of radial positions, directions, injected helicities, and shots. The normal-

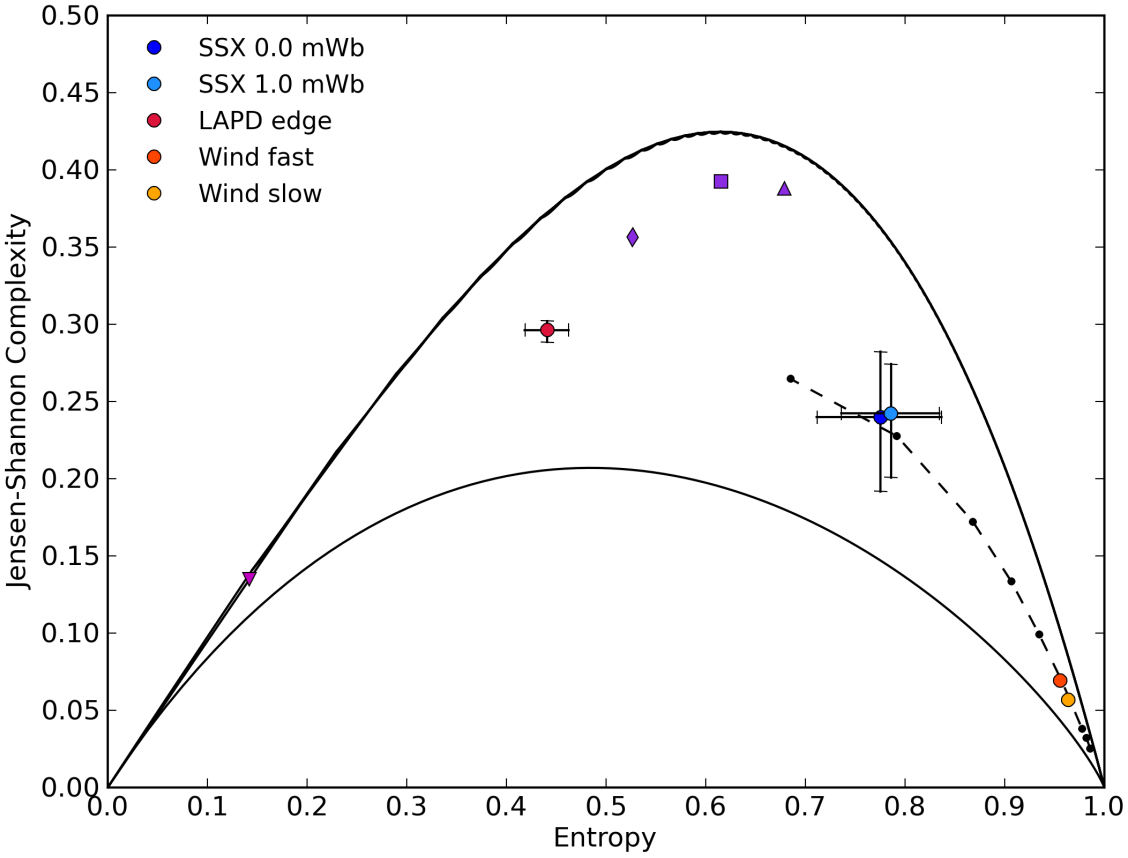

FIG. 1. The $n = 5$ CH plane with SSX $\dot{B}$ data for two injected helicities, *Wind* fast and slow stream B data, LAPD edge plasma ion saturation current signals, and paradigmatic chaotic, periodic, and stochastic systems for comparison. From left to right, the three purple markers represent chaotic skew tent, Henon, and logisitic maps. The sine function is shown in magenta, and stochastic fBm is shown in black for a range of Hurst exponents. Crescent shaped curves show the maximum and minimum possible C_{JS} for a given PE_{norm}.

ized permutation entropy and Jensen-Shannon comlexity were calculated for each series, using $n = 5$ in order to satisfy the common condition $L > 5n!$, as recommended in [1] and [18]. The length-preserving embedding delay method was employed to preserve this condition after sub-sampling. Each series was analysed from $40 - 60$ μs after discharge, as the turbulence signals exhibit a stationary phase during that period. SSX mangetic field fluctuation spectra resolve structures well only up to 9 or 10 MHz, so an embedding delay of $\tau = 8$ was used to limit the range of frequencies investigated while preserving enough statistics to avoid artificial numerical effects, which will be discussed later. The average position of series from all three directions of the inner 4 probe coils at each of two helicity settings is shown in blue in Figure 1. Error bars indicate standard deviations above and below the mean.

Figure 1 also shows the positions of both fast and slow stream magnetic fluctuations in the solar wind. The fast stream magnetic signal from *Wind* consisted of almost $230,000$ values, and the slow stream signal of over $170,000$. Since both signals were highly stationary, a set of subseries could be treated as an ensemble. The length of subseries was chosen in conjunction with the embedding delay so as to satisfy the condition $L_{wind}/\tau_{wind} = L_{ssx}/\tau_{ssx}$ while maximizing differentiation between the Jensen-Shannon complexities of the fast and slow streams and keeping the number of subseries on the order of 10. Based on these criteria, entropies and complexities were averaged over 20 subseries each $11,375$ values in length for the fast stream signal and 15 subseries of $11,375$ values for the slow stream. Delays of 70 were used, which limits the upper frequency range of the dynamics under investigation to well within the inertial range. Error bars are not displayed due to the small size of deviations in complexity and entropy from the mean.

Previous work using frequency spectra has suggested that the edge fluctuations of magnetized plasmas in the

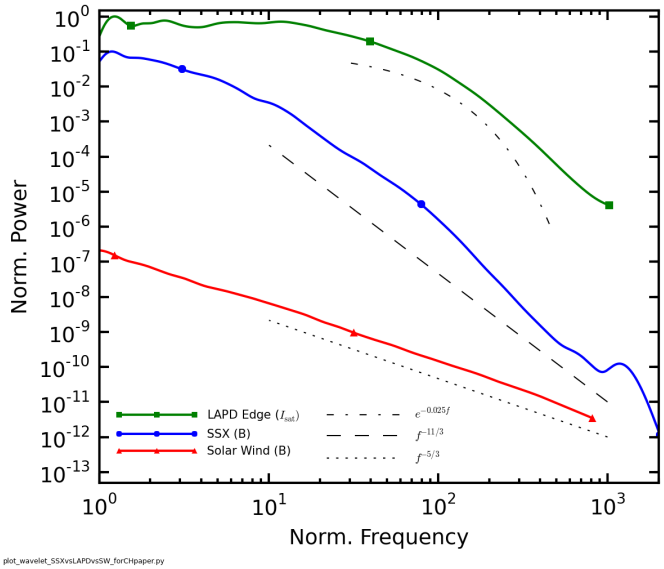

FIG. 2. Spectra of LAPD edge $I_{\rm sat}$ fluctuations, SSX magnetic fluctuations and Solar wind magnetic fluctuations. Each spectrum is normalized to a different time scale: LAPD is normalized to 500Hz; SSX is normalized to 12.8kHz; Solar wind is normalized to 165μHz. Analytic forms for an exponential, a -11/3 power-law and a -5/3 power-law (Kolmogorov) are indicated by the black lines. The power scale is arbitrary as the emphasis here is on the shape of the curves, not the relative power content of the spectra. The three spectra indicate a clear transition from exponential-like to power-law like broadband spectra.

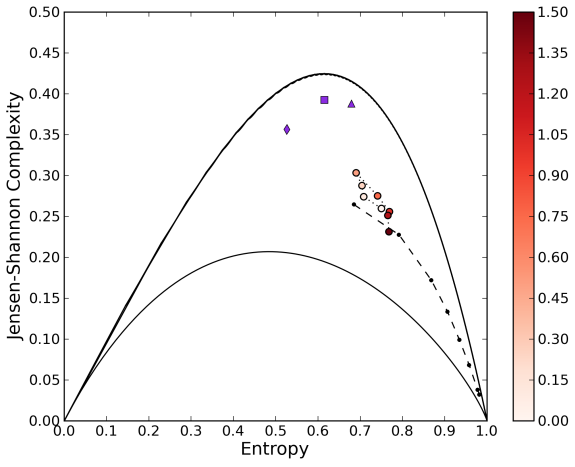

FIG. 3. Helicity scan of SSX $\dot{B}$ data. Dotted lines connect successive helicity settings to illustrate the trend. The color scale represents the stuffing flux setting, in mWb. Data was acquired from the innermost probe channel. Chaotic maps (purple) and fBm (dotted black curve) are shown for comparison.

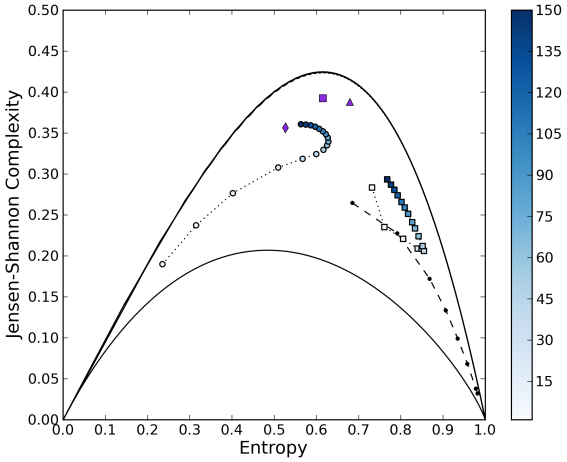

FIG. 4. Trajectories of SSX B and $\dot{B}$ data as functions of embedding delay from 1 to 150, color-coded by increasingly dark shades of blue. Model chaotic systems and stochastic processes are also shown.

Large Plasma Device (LAPD) and other devices are chaotic in nature [15]. For purposes of comparison, the permutation entropy and Jensen-Shannon complexity were also calculated for 1.5 MHz ion saturation current signals from a radial position corresponding to edge plasmas in the LAPD. The CH position of the LAPD edge plasma shown in Figure 1 in red was averaged over 25 shots and 5 sections of 1000 values for each shot with no embedding delay.

As expected, the LAPD edge plasma signals analyzed occupy the central region of the plane which has previously been associated with more chaotic than stochastic dynamics in experimental data [8, 16]. The SSX turbulence data at the 0.0 mWb and 1.0 mWb stuffing flux settings, on the other hand, occupy a higher entropy range comparable to fractional Brownian motion with positive correlation between successive increments of motion. The proximity of SSX $\dot{B}$ signals to stochastic processes like fBm suggest that these signals are dominated by stochastic rather than chaotic dynamics. Furthermore, the relative positions of SSX and LAPD sources on the $PE_{\rm norm} \times C_{JS}$ plane indicate that the SSX plasma source is more truly turbulent than the LAPD edge plasma.

According to these metrics, magnetic fluctuations in the solar wind near 1 AU more random than either wind tunnel MHD turbulence or even classical Brownian motion (fBm with Hurst exponent of 1/2). The fast stream signal exhibits slightly more entropy and less complexity than the slow stream signal, supporting identification of faster streams with more stochastic dynamics. Although it has been well documented that the solar wind exhibits well-developed turbulence (cite), this is the first time that developed MHD turbulence in an astrophysical plasma has been identified based on a complexity-entropy analy-

sis or compared in this manner to other plasma sources.

Although little spread was observed in either the *Wind* or LAPD data analyzed, significant spread was seen in the complexity and entropy of SSX $\dot{B}$ signals. Perhaps the most interesting contribution to this spread was from injected helicity. Figure 2 shows the helicity dependence of SSX $\dot{B}$ data, plotting CH positions for signals over a scan of stuffing fluxes measured at the innermost channel of the probe, averaged over directions and shots. The same ordinal pattern parameters used to generate Figure 1 ($n = 5$ and $\tau = 8$) were used in Figure 2. There is a clear increasing trend in complexity up to 0.5 mWb, after which the complexity decreases with stuffing flux. As a result the data executes a tight "loop" in the plane extending up just past $C_{JS} = 0.3$ with normalized entropies between 0.68 and 0.78. Thus the degree of twistedness in the injected spheromak has consequences for the degree of correlational structure in the the resulting relaxation dynamics.

In some cases, additional information can be gained about the underlying dynamics of a system by examining the permutation entropy and Jensen-Shannon complexity as functions of the embedding delay τ [26]. Figure 3 compares SSX $\dot{B}$ signals averaged over all shots, directions, and inner four probe channels to the corresponding B signals obtained by integrating $\dot{B}$ for a scan of the embedding delay from $\tau = 1$ to $\tau = 150$ ($\tau = 1$, 5, and then 10, 20, etc. in steps of 10). As before, the length-preserving method was employed. However, despite the use of the length preserving method, useful statistics seem only to be obtainable for our 1300 value series up to $\tau \approx 50$. For larger embedding delays, the number of length n segments used in calculating the ordinal pattern distribution is reduced by over 15%. The poorly "filled out" PDF manifests itself in artificially high complexities- note the abrupt hook shapes executed by the trajectories of both B and $\dot{B}$ data near $\tau = 50$.

IV. Conclusion

In this paper, fully dynamical laboratory and astrophysical plasmas have been studied using the ordinal pattern-based CH plane introduced by Rosso *et al* for the first time. Comparing the relative positions of drift-wave, MHD wind tunnel, and solar wind plasmas, it was found that the three systems occupy different regions of the CH plane, suggesting that despite the broad-band spectra exhibited by all these systems, the CH analysis is capable of clearly differentiating between drift-wave, partially developed, and fully developed turbulence. Considering that drift-wave turbulence is thought to be a result of the nonlinear interactions of relatively few modes while fully developed turbulence contains too many modes to distinguish, it would appear that the CH positions of these magnetized plasmas are reflective of the number of degrees of freedom of the system in question. In particular, the smaller number of modes generating drift-wave turbulence in LAPD edge plasmas may be reflected by the low-middle entropy and middle-range complexity of that system, while the high entropy and low complexity of magnetic fluctuations in the solar wind may reflect the multitude of degrees of freedom active in that system. Based on the relative CH positions of SSX MHD wind tunnel and *Wind* data, although SSX is on its way towards the highly stochastic turbulence in the solar wind, this analysis indicates that further steps are needed for SSX to most accurately model solar wind turbulence. The confined nature of the experiment and short lifetimes involved are both potential contributors to the discrepancy in CH positions. After all, besides the boundary conditions imposed by astrophysical bodies, the solar wind is an unconfined and extremely long lived plasma. Whether one or both of these parameters could be varied to reduce the complexity and increase the entropy of SSX to that of the solar wind is an open question. In any case, the CH methodology has provided us with one more means of comparison and a clear goal to work towards.

[1] J. M. Amigó, S. Zambrano, and M. A. F. Sanjuán. Combinatorial detection of determinism in noisy time series. *EPL (Europhysics Letters)*, 83(6):60005, 2008.

[2] Christoph Bandt. Ordinal time series analysis. *Ecological Modelling*, 182(34):229 – 238, 2005. Scaling, fractals and diversity in soils and ecohydrology.

[3] Christoph Bandt and Bernd Pompe. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.*, 88:174102, Apr 2002.

[4] Aurelio Fernandez Bariviera, Luciano Zunino, M Belón Guercio, Lisana B Martinez, and Osvaldo A Rosso. Efficiency and credit ratings: a permutation-information-theory analysis. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(08):P08007, 2013.

[5] Xavier Calbet and Ricardo López-Ruiz. Tendency towards maximum complexity in a nonequilibrium isolated system. *Phys. Rev. E*, 63:066116, May 2001.

[6] A. Wan M.R. Brown D.A. Schaffner, V.S. Lukin. Turbulence analysis of an experimental flux rope plasma. *Plasma Physics and Controlled Fusion*, 56:20, 2014.

[7] David P Feldman and James P Crutchfield. Measures of statistical complexity: Why? *Physics Letters A*, 238(45):244 – 252, 1998.

[8] Walter Gekelman, Bart Van Compernolle, Tim DeHaas, and Stephen Vincena. Chaos in magnetic flux ropes. *Plasma Physics and Controlled Fusion*, 56(6):064002,

2014.

[9] T. Gray, M. R. Brown, and D. Dandurand. Observation of a relaxed plasma state in a quasi-infinite cylinder. *Phys. Rev. Lett.*, 110:085002, Feb 2013.

[10] Denis Jordan, Gudrun Stockmanns, Eberhard F. Kochs, Stefanie Pilge, and Gerhard Schneider. Electroencephalographic order pattern analysis for the separation of consciousness and unconsciousness: An analysis of approximate entropy, permutation entropy, recurrence rate, and phase coupling of order recurrence plots. *Anesthesiology*, 109(6), 2008.

[11] A.M. Kowalski, M.T. Martín, A. Plastino, and O.A. Rosso. Bandt-pompe approach to the classical-quantum transition. *Physica D: Nonlinear Phenomena*, 233(1):21 – 31, 2007.

[12] P.W Lamberti, M.T Martin, A Plastino, and O.A Rosso. Intensive entropic non-triviality measure. *Physica A: Statistical Mechanics and its Applications*, 334(12):119 – 131, 2004.

[13] Duan Li, Xiaoli Li, Zhenhu Liang, Logan J Voss, and Jamie W Sleigh. Multiscale permutation entropy analysis of eeg recordings during sevoflurane anesthesia. *Journal of Neural Engineering*, 7(4):046010, 2010.

[14] R. López-Ruiz, H.L. Mancini, and X. Calbet. A statistical measure of complexity. *Physics Letters A*, 209(56):321 – 326, 1995.

[15] J E Maggs and G J Morales. Exponential power spectra, deterministic chaos and lorentzian pulses in plasma edge dynamics. *Plasma Physics and Controlled Fusion*, 54(12):124041, 2012.

[16] J E Maggs and G J Morales. Permutation entropy analysis of temperature fluctuations from a basic electron heat transport experiment. *Plasma Physics and Controlled Fusion*, 55(8):085015, 2013.

[17] E. Olofsen, J. W. Sleigh, and A. Dahan. Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect. *British Journal of Anaesthesia*, 101(6):810–821, 2008.

[18] M. Riedl, A. Müller, and N. Wessel. Practical considerations of permutation entropy. *The European Physical Journal Special Topics*, 222(2):249–262, 2013.

[19] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes. Distinguishing noise from chaos. *Phys. Rev. Lett.*, 99:154102, Oct 2007.

[20] Patricia M. Saco, Laura C. Carpi, Alejandra Figliola, Eduardo Serrano, and Osvaldo A. Rosso. Entropy analysis of the dynamics of el niño/southern oscillation during the holocene. *Physica A: Statistical Mechanics and its Applications*, 389(21):5022 – 5027, 2010.

[21] M.C. Soriano, L. Zunino, O.A. Rosso, Ingo Fischer, and C.R. Mirasso. Time scales of a chaotic semiconductor laser with optical feedback under the lens of a permutation information analysis. *Quantum Electronics, IEEE Journal of*, 47(2):252–261, Feb 2011.

[22] Xiaoliang Sun, Yong Zou, Victoria Nikiforova, Jurgen Kurths, and Dirk Walther. The complexity of gene expression dynamics revealed by permutation entropy. *BMC Bioinformatics*, 11(1):607, 2010.

[23] Vinita Suyal, Awadhesh Prasad, and HarinderP. Singh. Hysteresis in a solar activity cycle. *Solar Physics*, 276(1-2):407–414, 2012.

[24] Massimiliano Zanin, Luciano Zunino, Osvaldo A. Rosso, and David Papo. Permutation entropy and its main biomedical and econophysics applications: A review. *Entropy*, 14(8):1553–1577, 2012.

[25] X. Zhang, D. Dandurand, T. Gray, M. R. Brown, and V. S. Lukin. Calibrated cylindrical mach probe in a plasma wind tunnel. *Review of Scientific Instruments*, 82(3):–, 2011.

[26] L. Zunino, M. C. Soriano, and O. A. Rosso. Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach. *Phys. Rev. E*, 86:046210, Oct 2012.

[27] Luciano Zunino, Benjamin M. Tabak, Francesco Serinaldi, Massimiliano Zanin, Daro G. Prez, and Osvaldo A. Rosso. Commodity predictability analysis with a permutation information theory approach. *Physica A: Statistical Mechanics and its Applications*, 390(5):876 – 890, 2011.

[28] Luciano Zunino, Massimiliano Zanin, Benjamin M. Tabak, Daro G. Prez, and Osvaldo A. Rosso. Complexity-entropy causality plane: A useful approach to quantify the stock market inefficiency. *Physica A: Statistical Mechanics and its Applications*, 389(9):1891 – 1901, 2010.