
Report of Referee 1

This is an interesting manuscript which takes an approach to the problem of plasma turbulence which is very different from most analyses recently published. Three different turbulent plasmas, two in laboratories and the solar wind, are analyzed using the Bandt-Pompe permutation entropy and the Jensen-Shannon statistical complexity. The problem for me in reviewing this manuscript is that I am not at all familiar with either of these two formalisms, and the manuscript does not provide much help to me in increasing that familiarity.

The conclusions of this manuscript make sense; from very basic arguments I would agree that the unbounded solar wind is more likely to be in a high-entropy, fully developed turbulent state than laboratory plasma turbulence which is constrained by the walls of the confining vessels. But I am quite far from understanding the mathematical formalism which leads to the conclusions of this manuscript. So my request to the authors is not to add any more calculations to this manuscript, but to rewrite and re-express their presentation to make it more accessible to the large body of Physical Review readers who are not well-versed in the Bandt-Pompe formalism.

We thank the reviewer for this thoughtful suggestion and we've taken it to heart. The new manuscript features expanded text with some clarifying examples.

What is the physical interpretation of the statement that the solar wind "occupies the lower-right region of the CH plane"?

It means that virtually every one of the n! = 120 ordinal patterns in the solar wind magnetic field waveform was nearly *equally* represented. If each pattern was exactly equally represented, then the permutation entropy would be the sum from 1 to n! of -p log(p) = -(1/n!) log(1/n!) = log(n!). This is the maximal entropy, and subsequent permutation entropies are normalized to this value. The fact that the normalized permutation entropy of the solar wind is 0.95 says that every permutation was nearly equally represented. A more clear description of the nature of the time series based on what region of the CH plane they occupy has been added to the manuscript.

Both entropy and complexity are defined here. But, although entropy is a concept which most physicists are familiar with, I believe the same is not true of complexity. What's the physics of the statement that the laboratory turbulence examples considered here "have less permutation entropy and more statistical complexity"?

The interpretation of the Jensen-Shannon complexity is discussed in more detail in the revised manuscript, but a short summary is given here. Statistical complexity can be thought of as measuring the “non-triviality” of the distribution of ordinal patterns appearing in a time series. Systems whose dynamics we commonly think of as more structured have ordinal pattern distributions which are far from either the maximal order case of only one fluctuation ever occurring or the maximal randomness case of all possible fluctuations occurring. Therefore the complexity can also be seen as a measure of correlational structures in a system’s dynamics. The physical origins of these structures naturally vary widely from system to system. When applied to MHD turbulence, greater statistical complexity is thought to reflect a greater degree of structure in the magnetic field, although precisely what kind of structure remains to be seen in future work.

Figure 1 is hard to read; the symbols and labels are in very small print, the “downward pointing triangle” is not visible to my eyes (Is not the entropy of a precisely-defined sine wave zero?), and the “chaotic skew tent, Henon, and logistic maps” are not explained. Figure 2 is clearer, perhaps because it describes the well-known frequency spectra, but again many of the labels are very small and hard to read.

Figure 1 has been made larger and more readable.

If solar wind turbulence has less statistical complexity, why is it more “stochastic-like” (page 4 below Figure 2), and why is LAPD turbulence more “chaotic” (page 4, near the top the right-hand column).

The terms chaotic and stochastic are used in the mathematical sense, i.e. to refer to inherently probabilistic mathematical models and those which are deterministic yet highly unpredictable, respectively. Mathematical models of these two types robustly occupy distinct regions of the CH plane, with chaos having moderate entropy and high complexity and stochastic processes tending to have high entropies and low complexities. So when we refer to turbulent plasmas as stochastic-like versus chaotic-like, we are using an operational definition based on their proximity on the CH plane to mathematical systems of these two types. The logic to these operational definitions is that saying something is “chaotic-like”, for example, reflects deeper facts about the fluctuations accessible to that system more or less common to all and only “chaotic-like” systems.

In summary, I think that there is the potential for new insight from the unfamiliar formalism used for the analysis, and the conclusions make physical sense. But the authors need to rewrite the manuscript to clarify and illuminate (perhaps through simple examples) the meaning of the various terms used here as well as the overall physical concepts involved. I would be willing to review a revised version.

We thank the referee for their careful reading of our manuscript. In particular, the suggestion of illumination through simple examples has been taken to heart. We now feel it is ready for publication in PRE.

Report of Referee 2

I have carefully read the manuscript by Weck et al., submitted for publication on PRE.

The authors present the results obtained through the analysis of permutation entropy and complexity of three different plasma turbulence datasets. The results are very interesting, as they show the power of the data analysis technique quite convincingly and provide additional new knowledge on the systems studied here. I therefore warmly recommend their publication on PRE after a few points have been considered.

However, while I do not have any major remarks on the scientific content of the work, the manuscript certainly needs some text revision and reorganization before publication.

Here is a list of suggestions, comments and questions for the authors.

1. The Introduction section is oddly organized. After a quick introduction to the general framework, the authors describe the main results of the paper, before defining the various actors of their work. For example, they write of LAPD and SSX without having introduced the experiments. The whole second paragraph ("We compute the values...") might be irrelevant at this point, and perhaps moved forward in the paper, or, at least, made more general.

We have expanded the section on the background of the datasets used, as well as have added some brief introductory phrases regarding the datasets near the beginning of the introduction. However, we believe the emphasis of this work to be on the analysis technique and the computed values of Permutation Entropy and Jenson-Shannon Complexity rather than on particular results from the experiments or observations so our the introduction focuses on first relaying the overall results of the analysis, before moving on to describe the data in more detail.

2. Still in the introduction, on page 2, the three datasets should be described in a dedicated section. Also, I'd rather have a different sorting of the experiments, leaving the solar wind for last (it's the only natural plasma, and it doesn't make saense to me to describe it in between two laboratory devices). Also, the paragraph describing LAPD is somewhat unbalanced with respect to Wind and SSX. I frankly do not know the device, and the text doesn't help understanding what kind of experiment it is. I'd suggest the authors to spend more words in the description of the LAPD.

The detailed description of the devices and datasets has been reorganized to discuss the experiments first, then the solar wind. The description of the LAPD has been expanded as well.

3. Section II: although the definition of the quantities used here (PE and C) is rigorous, I think this section lacks a short intuitive description of their meaning. Could the author explain better in intuitive way the "information" meaning of the two quantities, and how they should be interpreted?

A more detailed background and intuitive description of these quantities has been included.

4. Page 3, 10 lines after eq.n (3): The authors describe the CH plane without properly introducing it. I would suggest to add a brief description of "what is the CH plane", and how it should be read (complexity vs entropy: what does it says to the reader?). Also, figure 1 is not easy to read especially if printed in BW. The markers are not easy to distinguish, so I'd suggest the authors to separate the marker style more efficiently. For example, the values for the maps could be indicated using open markers, which would make them more easily recognizable with respect to the experimental ones.

The term "CH plane" has been better described. The figure has also been updated for readability and clarity.

5. SECTION III: Is it possible that integration of B introduces some artifact (correlations) on the SSX time series? While it is expected that the two time series will give different results, I understand this is unavoidable, but perhaps the authors could cross-check the "direction" of the changes by performing a complementary operation on the Wind data, i.e. by estimating the dB/dt for that dataset?

The effect of integration could potential introduce some artifacts at the level of one, two or three time steps (the integration procedure used is trapezoid rule). However, the effect should be diminished by the time larger time-steps are used. For the values presented here, the integration should not be a significant factor given the time-step used.

6. page 3, line 15 from the bottom: the description of the plot for SSX data is not very easy to read. For example, the author never mentioned the existence of 4 probes (they only talk about an array of probes in the introduction). Also, "position" and "direction" are too generically used in the same sentence, so they are a bit confusing. Perhaps "position in the CH plane" and "component" (rather dan direction) could make the reading smoother?

"The inner four probes" references the four innermost loops of the 16 channel magnetic probe array referenced in the discussion of the SSX device. The sentence has been clarified to reflect that the single point shown in the figure (for each helicity setting) consists of the average computed values for the individual orthogonal direction of each of the four inner most probe locations of the probe array.

7. page 4,figure 2: the exponents of the reference power-laws are too small and difficult to read.

We have modified the figures so they are readable at standard magnification.

8. page 4, second paragraph ("Previous work using..."): is it really OK to compare analysis of different systems done with or without using the delay?

Since there is, in general, no preferred choice of embedding delay, there should be no reason that datasets of different embedding delays cannot be directly compared as long as the embedding delay is made small enough to avoid a lack of statistics which can introduce numerical artifacts. For this analysis, the embedding delays were chosen for three main reasons: 1) To minimize any numerical artifacts due to time series length 2) To avoid contamination from uncorrelated noise and 3) To include the speculated physical mechanism contributing to the nature of the fluctuation. For the solar wind, a larger time delay was chosen to minimize the contribution of dissipative effects and focus primarily on inertial range turbulence. For SSX, a high frequency noise inherent to the discharge is avoided by a larger time delay. For LAPD, as the previous work phrase indicates, the fluctuations are generated by non-linear interactions of modes which generate Lorentzian structures in the signal. Unfortunately, the data was sampled at a rate that put the dominant temporal scale of these structures at the maximum sampling frequency. Thus no embedding delay was used. If a larger delay was used, the physics of interest in the LAPD is lost. Ideally, data taken at a higher cadence would be used (and this will be part of future work) so that the embedding delay can be pushed higher, but we are confident this will ultimately have no effect on the results. The LAPD data does not suffer from the high frequency noise issue as SSX does.

9. Page 4, in the description of the solar wind results, it is odd that the fBm is less stochastic than the wind. Perhaps this point deserves some more discussion.

As mentioned in the text, the fBm curve shown was generated by varying the correlation between increments in the fBm model from a negative correlation to a positive one. fBm whose Increments have zero correlation (Hurst exponent = ½) is identified with Brownian motion. As noted, the CH position of a time series from Hurst exponent = ½ fBm was found to be less entropic and more complex than the solar wind time series analyzed. (although fBm with large negative correlations between increments remains more stochastic than the solar wind). Exactly how this result is to be interpreted is not clear, although perhaps by comparing correlations between increments in the solar wind to those in the fBm model a more detailed understanding could be developed. Such connections between stochastic models such as fBm and turbulence in the solar wind based on the CH methodology may be explored in future work.

10. In the same discussion: the higher level of stochasticity of fast solar wind is consistent with the presence of alfvénic fluctuations (which are more uncorrelated than the turbulent *fluctuations), and with the generally observed higher degree of intermittency of slow wind (suggesting a more developed turbulence).*

This is a good point. Although we haven't checked Alfvenicity for this data set (we only have access to B), we will add a statement and a reference to the manuscript hypothesizing this possibility. In truth, the fast and slow samples of Wind data both have H very nearly unity and C nearly zero, so it's difficult to make strong claims corresponding to fast versus slow wind.

11. page 4, discussion on SSX results: once again, it is to be expected that dB/dt is more complex than B. This is probably even more the case for intermittent turbulence. The difference between the "slopes" (-7/3 and -11/3) is compatible with the extra frequency coming from the derivation of B (which would give an -8/3 slope, assuming that E(B)=f^-11/3). The authors could comment on that.

We appreciate the reviewer for pointing this out. We will consider this in future work. However, the specific slope lines indicated in the plot are there mainly for reference and not as an indication of any theoretical model for the data.

12. The discussion about the shape of the power spectra is not very clear, and perhaps should be rephrased. In my opinion, this part of the analysis is also not fully objective, and might need some correction. From a visual inspection, it is not very easy to support the authors' claims: while it is evident that the solar wind magnetic spectrum is (as well known) a power law with Kolmogorov-type exponent, and while the LAPD spectrum has a somewhat exponential shape (despite its tail is clearly shallower than exponential, and the comparison with exponential curve in the Figure highlights the discrepancy), I do not clearly see the power-law-ish behavior of the SSX spectra, which in my opinion have a very similar shape as for LAPD.

I still agree with the interpretation given by the authors, but I find it only strictly valid for Wind and LAPD, while the SSX cases stays somewhere in between (as claimed), but not necessarily because a power-law is visible. Perhaps the authors should find a more precise representation for these curves, or at least present a more neutral comparison with reference curves (e.g.: compare with both power-law and exponential for all four cases). Or, at the very least, stress the approximation of the observation.

The power-law nature of the SSX data is discussed in previous work (Schaffner PPCF 2014 and Schaffner ApJ 2014). However, we can certainly see the reviewer's point about the shape of the SSX curve. In the above cited papers, the SSX spectra is shown to have two different power-law regions, roughly surround the 1MHz scale which is thought to be at a dissipation scale in the SSX plasma. This can be seen as the inflection point in the middle of the SSX spectra in the manuscript. The transition region between these two power-law regions can in fact be modeled by an exponential fit. Such an exponential modification at dissipation scales in currently being explored (see Terry Phys Plasma 2012). Thus, the SSX data can be viewed as having a combination of power-law like scaling and exponential like scale which places it in

an intermediate regime. The text has been updated to emphasize this point. We considered attempting to demonstrate this point graphically in figure 2, but decided that attempting to include multiple power-law reference lines and exponential reference line to the SSX data would have made things to cluttered and overly complicated. We have restricted the horizontal range to emphasis the spectral region of interest. The primary point is that SSX data is generally in between the other two data sets.

13. Page 5, second paragraph: just as a hint, it could be interesting to check what happens to more turbulent laboratory plasmas, as for example in RFP devices (the Italian RFX is a good example of fully developed turbulence with intermittency, see several works by Carbone, Antoni, Martines and co-authors).

We don't have access to the RFX data described by the reviewer but we do have access to liquid metal turbulence data from the Madison Dynamo Experiment. This data occupies a region of the plan comparable to the solar wind, except at particular time scales at which it temporarily increases in complexity and decreases in entropy, perhaps indicating a scale with greater physical structure. We plan to continue this work with comparisons to other MHD-type flows, including reversed field pinches. We may be able to obtain data from the Madison MST reversed field pinch device.

14. Finally, again just a suggestion, it would be very interesting to explore the role of high order moments (intermittency) besides the power spectra. In fact, since intermittency plays a relevant role in the build-up of small scale correlations, this could explain (even more than the spectral shape and slope) the differences between all experiments. It is possible that the laboratory plasmas do not show clear intermittency, because of their low level of turbulence (Re=100), but this could be interesting at least for the fast/slow wind.

In fact the SSX fluctuations do exhibit intermittency (see Schaffner PPCF 2014 and Schaffner PRL 2014) and appear to have similar intermittent characteristics as the solar wind. LAPD data has also been shown to exhibit intermittency (Carter Phys Plasmas 2006). However, the reviewer is right in that a comparison amongst these datasets from the point of view of intermittency and structure functions would be of definite interest for future work.

We thank the referee for their careful reading of our manuscript. We have adopted most of the suggestions. We now feel it is ready for publication in PRE.