Response to the referees and summary of changes

Date: March 11, 2019

Dear editor,

We would like to resubmit our revised manuscript entitled

Ordinal partition transition network based complexity measures for inferring coupling direction and delay from time series

by Y. J. Ruan, R.V. Donner, S. G. Guan, and Y. Zou,

originally submitted for publication to *Chaos* with the manuscript number **#181384**.

Thank you very much for the comments and suggestions of the two referees. All these comments helped us to improve the manuscript. We have marked all changes in the manuscript in red color to facilitate the reading.

Please find below a detailed response to all specific remarks and questions raised by the referees. We hope that our revised manuscript is now acceptable for publication in *Chaos*.

Sincerely yours,

Y. J. Ruan, R.V. Donner, S. G. Guan, and Y. Zou

Response to the first referee

The authors are claiming to have proposed three measures for inferring coupling directions and delays from time series based on ordinal patterns. The first one, based on the standard deviation for conditional probabilities, might be new. But, the second one and the third one have been defined previously as far as I searched: The second one seems to be used at least in Li and Ouyang, NeuroImage 52, 497 (2010), while the third one was used in Staniek and Lehnertz, Phys. Rev. Lett. 100 158101 (2008). Thus, the novelty for the manuscript should be compromised and the authors should cite these preceding papers and allocating their contribution in the existing literature more appropriately. Their results shown in the figures look promising. Especially, the coupling direction and delay for the daily surface air temperature record from Oxford to that of Vienna sound reasonable. Therefore, the authors should give the credits for the related previous papers properly for the publication in Chaos.

We thank the referee for pointing these references, which have been properly cited in the revision. We agree that our measures have much similarities with those that have been pointed by the referee, but by the same time, we wish to emphasize the difference that has been included in the revision:

The second complexity measure $H_{X \to Y}(\tau)$ and the third complexity measure KLD indeed shows much similarity with the method proposed by Li et al, and respectively, the symbolic transfer entropy as proposed by Staniek and Lehnertz. However, the conceptional difference is the following. The time delayed conditional mutual information quantifies the interaction information from X to Y by the difference between $p(\pi_j^Y(\tau); \pi_j^Y | \pi_i^X)$ and $p(\pi_j^Y(\tau); \pi_j^Y)$. On the other hand, the transfer entropy characterizes the coupling by the amount of uncertainty reduced in future values of $\pi_j^Y(\tau)$ knowing the past values of π_i^Y only if π_i^X would be the possible causal effect. More specifically, the information transfer from X to Y is computed by the divergence between two probabilities of $p(\pi_j^Y(\tau) | \pi_j^Y, \pi_i^X)$ and $p(\pi_j^Y(\tau) | \pi_j^Y)$.

Furthermore, there is much difference with the choice of parameter τ. In the reference of Li et al, τ is suggested not to be less than the embedding dimension D following the traditional way of computing conditional mutual information. On the other hand, in the reference of Staniek et al, τ is treated as an algorithmic parameter which had been suggested as $\tau = 1$. Furthermore, it has been assumed that the empirical choice of $\tau = 1$ is always applicable if a meaningful natural sampling rate of time series is a available. However, our method interprets τ differently, considering to extract the proper coupling delay τ explicitly by treating it as a dynamical unknown variable.

We have included the discussion in the revision when introducing the complexity measure.

In the abstract and the main text, OPTN is defined as an abbreviation for two different phrases. In the abstract, it comes from ordinal pattern transition networks, while in the main text, it comes from ordinal partition transition networks. Please define the abbreviation by one of them.

In the revision, we use of the acronym for OPTN as ordinal partition transition network, which keeps the consistency with the literature. This has been corrected for both the abstract and the main text.

Response to the second referee

The main purpose of the manuscript is to show that ordinal pattern transition networks (OPTNs) from time series can provide an approach to improve our understanding of the underlying dynamical system (responsible for the data). The authors conclude that OPTNs can be used as complementary tools for causal inference tasks.

The text is a well-written and scientifically sound manuscript that adds understanding to the important area of time series analysis. Before its publication, I believe some questions must be better answered by the authors.

We thank the referee for the overall positive evaluation on our manuscript.

1- On figures 4 - 8 and 11. The authors claim the results are "based on averages over 20 independent realizations of the considered processes". So the dispersion of the mean values (plotted by error bar or any other way) must be included. A discussion about dispersion values will bring more information about the validity of the results.
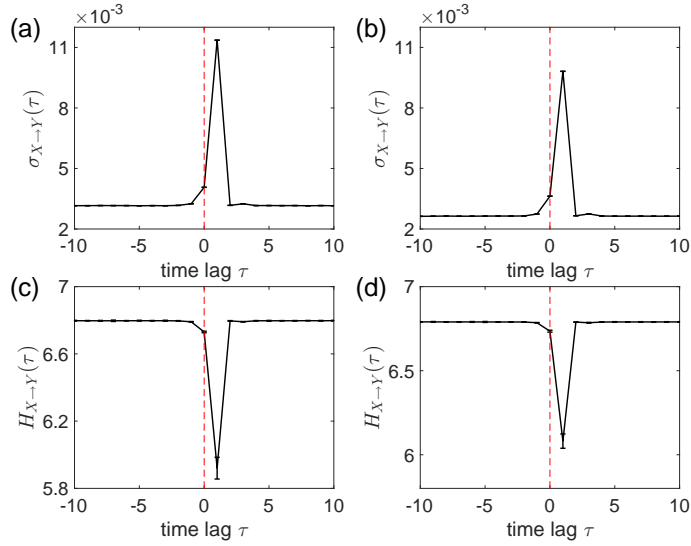
FIG. 1: (Color online) Values of three ordinal pattern co-occurrence complexity measures in dependence on the mutual delay τ for realizations of Eq. (1): (a,b) $\sigma_{X \to Y}(\tau)$, (c,d) $H_{X \to Y}(\tau)$, (e,f) KLD(τ). (a,c,e) show the results for realizations from the linear model, while (b,d,f) correspond to the nonlinear transformation of the first subsystem as described in the text. The vertical red dashed lines indicate the values for vanishing mutual delay ($\tau = 0$). Note that the error bars have been enlarged by 1000 times.

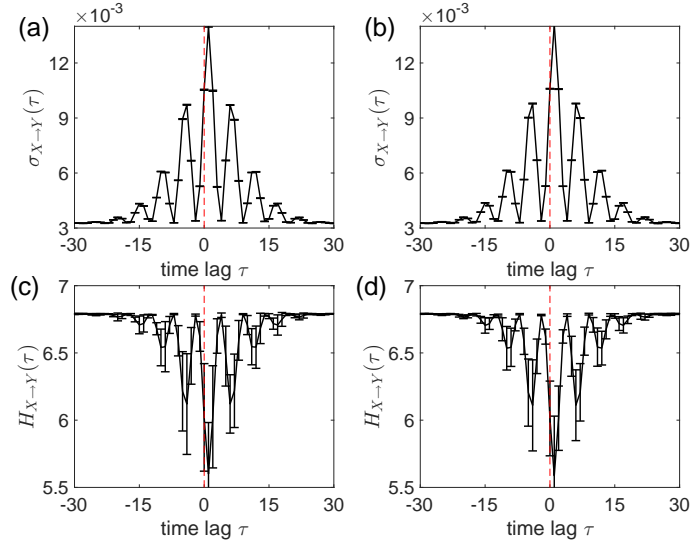FIG. 2: (Color online) Same as in Fig. 1 but for Eq. (10). Note that the error bars have been enlarged by 1000 times.

This is indeed a nice comment. In the very beginning, we plan to include the error bars in all figures. However, we show in this reply letter, these error bars are really invisible even though we have magnified by 1000 times as shown in Figs. 1 - 4. Later we decided not to include the error bars in the manuscript since they do not provide additional information. We have clarified the details for the numerical results and suggest not to include the plots with error bars.

2- Figures 9 and 10 show results only for KLD measure. Why not for all quantifiers? Figures 9 and 10 also show that KLD is a reliable quantifier only for time series larger than 10000. Such a result is consistent for both systems considered in the figures. Fig. 11 shows results for an experimental situation where the available data size is on the same order (10000). So the results must be corroborated at least using other quantifiers. A sentence discussing it will be also appreciated.
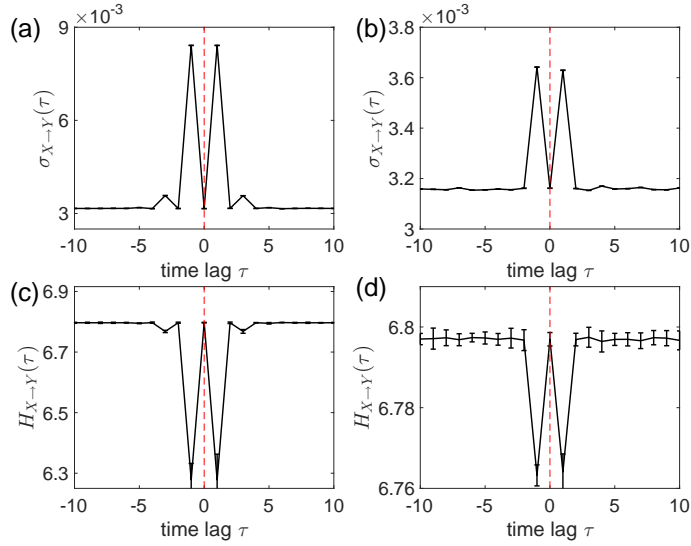
FIG. 3: (Color online) Same as in Fig. 1 but for Eq. (11). Note that the error bars have been enlarged by 1000 times.
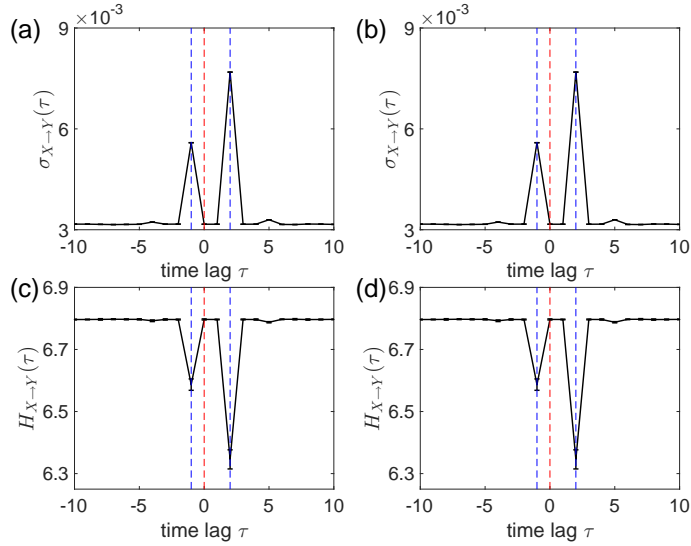

FIG. 4: (Color online) Same as in Fig. 3 but for Eq. (12). The two employed delays of $\tau = -1$ and $\tau = 2$ are additionally highlighted by vertical blue dashed lines. Note that the error bars have been enlarged by 1000 times.

We have clarified the motivations of KLD measures in the main body of the manuscript in the revision. The main reason is that the KLD measure is more statistically motivated which may not be subjected to the marginal distributions of ordinal patterns. This is expected from the computations of probabilities. However, the other two of $\sigma_{X \to Y}$ and $H_{X \to Y}$ are empirical.

Anyway, from the viewpoint of numerical studies, We have obtained consistent results for all measures of KLD and $\sigma_{X \to Y}$ and co-occurrence complexity entropy $H_{X \to Y}$. In a word, all measures show asymptotical convergences to zero values for uncoupled systems, while positive non-zero values.

We note that, together with the revision, we provide supplementary material file containing all results of finite size effects. Rather similar results have been obtained.

We have included the discussion as suggested, in particular the results are reliable if time series length larger than 10000.

1- Please, provide References for the sentences:

" In general, temperature records like those used here are characterized by some irregular alternation between colder and warmer phases due to large-scale atmospheric patterns (corresponding to high and low pressure centers, respectively), which in the case of western-to- central Europe travel in the majority of situations in eastward direction."

"This coupling delay is reasonable since temperature variations in Oxford arising from general atmospheric circulation patterns are typically propagated eastward and reach Vienna about 1to 2 days later, which corresponds to the average propagation speed of frontal systems in the northern mid-latitudes west-wind zone."

In the revision, we have included the proper reference when discussing this coupling delay, which can be attributed to teleconnections that two geological distant places are interacted via atmospheric circulations.