SUPPLEMENTAL MATERIAL FOR THE LETTER
Multiscale atmospheric dynamics: Cross-frequency phase–amplitude coupling in the air temperature

Milan Paluš*
Department of Nonlinear Dynamics and Complex Systems,
Institute of Computer Science, Academy of Sciences of the Czech Republic,
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic
(Dated: November 18, 2013)

In this Supplemental material we bring more information about the statistical evaluation of the conditional mutual information (CMI) as a measure of directional interactions, then a discussion of the frequency localization ability of the complex continuous wavelet transform (CCWT). Searching for clues to understanding of the physical mechanisms underlying the observed cross-scale interactions we point to some interesting facts about the North Atlantic Oscillation and its influence on the atmospheric circulation. However, for convenience of readers, we start with a brief introduction of the mutual information, the CMI in a general form and the form of the CMI used for the inference of causal relations from time series. We also make a note on the dependence of the CMI on the forward time lag and present examples of this dependence in the studied phase-amplitude interactions in the air temperature.

I. DEFINITION OF MUTUAL INFORMATION

Consider a discrete random variable X with a set of values Ξ. The probability distribution function (PDF) for X is $p(x) = \Pr\{X = x\}$, $x \in \Xi$. We denote the PDF by $p(x)$, rather than $p_X(x)$, for convenience. Analogously, in the case of two discrete random variables X and Y with the sets of values Ξ and Υ, respectively, their probability distribution functions will be denoted as $p(x)$, $p(y)$ and their joint PDF as $p(x, y)$. The *entropy* $H(X)$ of a single variable, say X, is defined as

$$H(X) = -\sum_{x\in\Xi} p(x)\log p(x), \tag{1}$$

and the *joint entropy* $H(X, Y)$ of X and Y is

$$H(X,Y) = -\sum_{x\in\Xi}\sum_{y\in\Upsilon} p(x,y)\log p(x,y). \tag{2}$$

The *conditional entropy* $H(Y|X)$ of Y given X is

$$H(Y|X) = -\sum_{x\in\Xi}\sum_{y\in\Upsilon} p(x,y)\log p(y|x). \tag{3}$$

The average amount of common information, contained in the variables X and Y, is quantified by the *mutual information* $I(X;Y)$, defined as

$$I(X;Y) = H(X) + H(Y) - H(X,Y). \tag{4}$$

The conditional mutual information $I(X;Y|Z)$ of the variables X, Y given the variable Z is given as

$$I(X;Y|Z) = H(X|Z) + H(Y|Z) - H(X,Y|Z). \tag{5}$$

For Z independent of X and Y we have

$$I(X;Y|Z) = I(X;Y). \tag{6}$$

By a simple manipulation we obtain

$$I(X;Y|Z) = I(X;Y;Z) - I(X;Z) - I(Y;Z). \tag{7}$$

Thus the conditional mutual information $I(X;Y|Z)$ characterizes the "net" dependence between X and Y without a possible influence of another variable, Z.

Consider now n discrete random variables $X_1, \ldots, X_n$ with values $(x_1, \ldots, x_n) \in \Xi_1 \times \cdots \times \Xi_n$, with PDF's $p(x_i)$ for individual variables X_i and the joint distribution $p(x_1, \ldots, x_n)$. The mutual information $I(X_1; X_2; \ldots; X_n)$, quantifying the common information in the n variables $X_1, \ldots, X_n$ can be defined as

$$I(X_1; X_2; \ldots; X_n) = \tag{8}$$

$$H(X_1) + H(X_2) + \cdots + H(X_n) - H(X_1, X_2, \ldots, X_n).$$

It is possible, however, to define mutual information functionals quantifying common information of groups of variables and also various multivariate generalizations of the conditional mutual information, see Ref. [1].

All the information theoretic functionals can be defined for continuous random variables. The sums are substituted by integrals and the PDF's by the probability distribution densities [2, 3]. Among the continuous probability distributions a special role is played by the Gaussian distribution. Let $X_1, \ldots, X_n$ be an n-dimensional normally distributed random variable with a zero mean and a covariance matrix $\mathbf{C}$. Then (see Refs. [1, 3] and references therein)

$$I_G(X_1; \ldots; X_n) = \frac{1}{2}\sum_{i=1}^{n}\log(c_{ii}) \; - \; \frac{1}{2}\sum_{i=1}^{n}\log(\sigma_i), \tag{9}$$

* mp@cs.cas.cz; http://www.cs.cas.cz/mp/

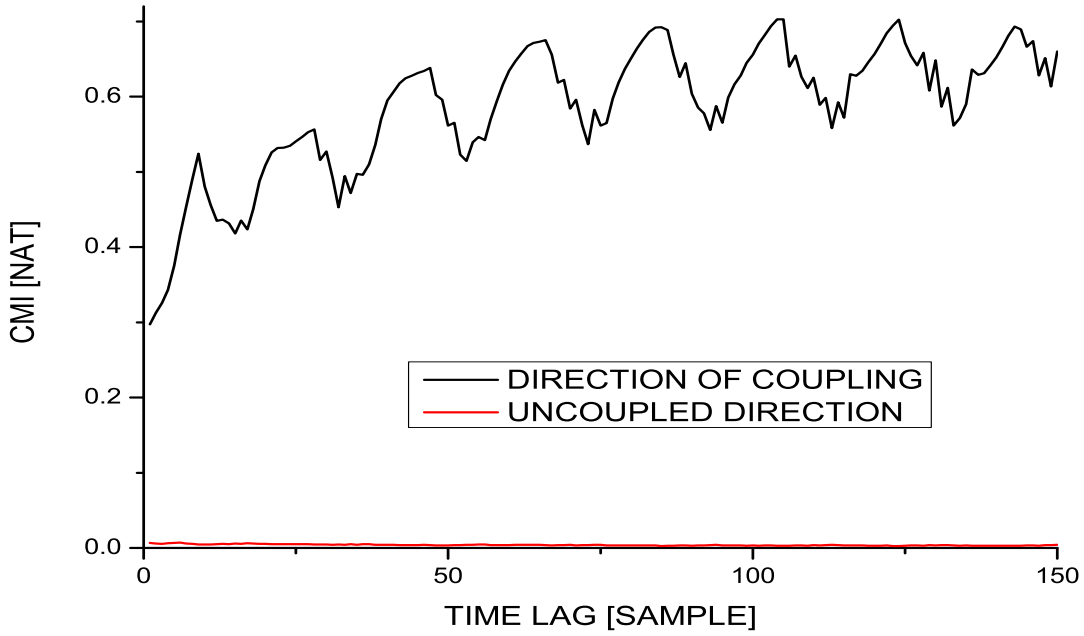

FIG. 1. The conditional mutual information $I(\phi_1(t);\phi_2(t+\tau)|\phi_2(t))$, characterizing the causal influence of the phase $\phi_1(t)$ of the autonomous Rössler system on the phase $\phi_2(t+\tau)$ of the driven Rössler system (the black curve); and the conditional mutual information $I(\phi_2(t);\phi_1(t+\tau)|\phi_1(t))$, characterizing the causal influence in the opposite direction $\phi_2(t) \to \phi_1(t+\tau)$ without any coupling (the red curve) as the function of the time lag τ.

where c_{ii} are the diagonal elements (variances) and σ_i are the eigenvalues of the $n \times n$ covariance matrix $\mathbf{C}$.

The entropy and information are usually measured in bits if the base of the logarithms in their definitions is 2, here we use the natural logarithm and therefore the units are called nats.

II. INFERENCE OF CAUSALITY WITH THE CONDITIONAL MUTUAL INFORMATION

Let $\{x(t)\}$ and $\{y(t)\}$ be time series considered as realizations of stationary and ergodic stochastic processes $\{X(t)\}$ and $\{Y(t)\}$, respectively, $t = 1, 2, 3, \ldots$. In the following we will mark $x(t)$ as x and $x(t+\tau)$ as x_τ, and the same notation holds for the series $\{y(t)\}$.

The mutual information $I(y; x_\tau)$ measures the average amount of information contained in the process $\{Y\}$ about the process $\{X\}$ in its future τ time units ahead (τ-future thereafter). This measure, however, could also contain an information about the τ-future of the process $\{X\}$ contained in this process itself, if the processes $\{X\}$ and $\{Y\}$ are not independent, i.e., if $I(x; y) > 0$. In order to obtain the "net" information about the τ-future of the process $\{X\}$ contained in the process $\{Y\}$ we use the conditional mutual information $I(y; x_\tau|x)$. The latter was used by Paluš et al. [4] to define the coarse-grained transinformation rate, able to detect direction of coupling of unidirectionally coupled dynamical systems.

We used the standard statistical language in which we considered the time series $\{x(t)\}$ and $\{y(t)\}$ as realizations of stochastic processes $\{X(t)\}$ and $\{Y(t)\}$, respectively. If the processes $\{X(t)\}$ and $\{Y(t)\}$ are substituted by dynamical systems evolving in measurable spaces of dimensions m and n, respectively, the variables x and y in $I(y; x_\tau|x)$ and $I(x; y_\tau|y)$ should be considered as $n-$ and $m-$dimensional vectors. In experimental practice, however, usually only one observable is recorded for each system. Therefore, instead of the original components of the vectors $\vec{X}(t)$ and $\vec{Y}(t)$, the time delay embedding vectors according to Takens [5] are used. Then, back in the time-series representation, we have

$$I\big(\vec{Y}(t); \vec{X}(t+\tau)|\vec{X}(t)\big) = I\Big(\big(y(t), y(t-\rho), \ldots, y(t-(m-1)\rho)\big); x(t+\tau)|\big(x(t), x(t-\eta), \ldots, x(t-(n-1)\eta)\big)\Big), \quad (10)$$

where η and ρ are time lags used for the embedding of the trajectories $\vec{X}(t)$ and $\vec{Y}(t)$, respectively. Only information about one component $x(t+\tau)$ in the τ-future of the system $\{X\}$ is used for simplicity. The CMI characterizing the influence in the opposite direction $I\big(\vec{X}(t);\vec{Y}(t+\tau)|\vec{Y}(t)\big)$ is defined in full analogy. Exactly the same formulation can be used for Markov processes of finite orders m and n. Based on the idea of finite-order Markov processes, Schreiber [6] has proposed a "transfer entropy" which is an equivalent expression for the conditional mutual information (10) – see Refs. [7, 8].

Paluš & Vejmelka [8] apply the CMI (10) in the inference of coupling directionality from time series generated by the unidirectionally coupled Rössler systems given by the equations

$$\begin{aligned} \dot{x}_1 &= -\omega_1 x_2 - x_3 \\ \dot{x}_2 &= \omega_1 x_1 + a_1\, x_2 \\ \dot{x}_3 &= b_1 + x_3(x_1 - c_1) \end{aligned} \tag{11}$$

for the autonomous system, and

$$\begin{aligned} \dot{y}_1 &= -\omega_2 y_2 - y_3 + \epsilon(x_1 - y_1) \\ \dot{y}_2 &= \omega_2 y_1 + a_2\, y_2 \\ \dot{y}_3 &= b_2 + y_3(y_1 - c_2) \end{aligned} \tag{12}$$

for the response system. Here we will use the parameters $a_1 = a_2 = 0.15$, $b_1 = b_2 = 0.2$, $c_1 = c_2 = 10.0$, and frequencies $\omega_1 = 1.015$ and $\omega_2 = 0.985$, and the coupling strength $\epsilon = 0.08$.

The CMI (10) can be evaluated either using the time series of the components x_1, y_1; or the phases ϕ_1, ϕ_2, obtained from the components x_1, y_1, respectively, using the analytic signal approach introduced in Eqs. (1)–(3) in the Letter. The CMI's $I(\phi_1(t);\phi_2(t+\tau)|\phi_2(t))$ and $I(\phi_2(t);\phi_1(t+\tau)|\phi_1(t))$ can be evaluated for a range of the forward time lag τ, see Fig. 1. Note that the τ-dependence of the CMI's reflects an oscillatory character of the underlying dynamics, rather than a time delay between the driving system and the driven system. In the example of the Rössler systems (11), (12) there is an "immediate" (zero-lag) coupling given by the diffusive term $\epsilon(x_1 - y_1)$ in (12). The study [8] suggests the use of the lag-averaged CMI as a more robust measure for the inference of causality than any single-lag CMI value, including the maximum CMI value over a range of the lags.

In the present study we use the CMI in order to infer directional, mixed phase-amplitude interactions. The functional

$$I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta)) \tag{13}$$

is evaluated and averaged for the forward lags τ from 1 to 750 days. This range was empirically established. We have studied the dependence of (13) on τ for different pairs of the time scales, i.e. pairs of the periods p_1, p_2 (p_1 for the phase ϕ_1 and p_2 for the amplitude A_2) – see the examples in Figs. 2–4. The range 1 – 750 days includes all observed patterns of positive CMI. Only the lag-averaged value of the CMI is used in the statistical testing in order to infer whether a directional influence exists or not. The time-lag dependence of the CMI is not a subject of this study.

III. STATISTICAL EVALUATION

Paluš & Vejmelka [8] also discuss numerical and statistical problems of inference of directional interactions from time series. They demonstrate how properties of dynamics underlying analyzed time series can influence estimators of conditional mutual information (or other measures of directional interactions). In particular, different levels of complexity and/or different typical time scales (main frequencies of oscillatory processes) can induce bias, e.g. estimates of the CMI quantifying interactions in the direction from slower to faster systems are greater than estimates of the CMI in the opposite direction without underlying causal interactions, i.e. either in the case of a symmetric coupling, or no coupling at all. Therefore the absolute values of CMI estimates are not informative, but it is necessary to relate the CMI values obtained from studied data to ranges of CMI values obtained from uncoupled processes which share important properties of analyzed data. This is the base of the surrogate data testing procedure in which we manipulate the original data in a randomization procedure which preserves original frequency spectra or variance on all relevant time scales; and the autocorrelation function or both the autocorrelation function and auto-mutual information function in the cases of the Fourier transform (FT) and the multifractal (MF) surrogate data, respectively. By construction, in the surrogate data the interactions between different time scales do not exist (FT surrogates), or only those explained by random cascades on wavelet dyadic trees are allowed (MF surrogates).

The specificity of causality inference methods can be secured by an appropriate choice of surrogate data which reproduce the bias of a CMI estimator applied to the studied data. The sensitivity of the method can be guaranteed by a certain bound on the estimator variance [8]. Typically, a lower variance requires a larger amount of data or, more precisely, a time series representing a longer epoch of evolution of the studied process [8, 9]. In an extensive numerical study we have found that the time series length 32768 daily SAT samples is necessary for reliable inference of the directed phase–amplitude interactions for the CMI functional (13) even though we use the estimator of mutual information (9) derived for Gaussian processes [1, 3, 10]. This type of the CMI estimator is also used by Molini et al. [11] in order to infer causality across rainfall time scales. However, here we specifically evaluate the influence of the phase ϕ_1 of slow oscillations on the amplitude A_2 of a higher-frequency variability.

Using the analytic signal approach introduced in Eqs.

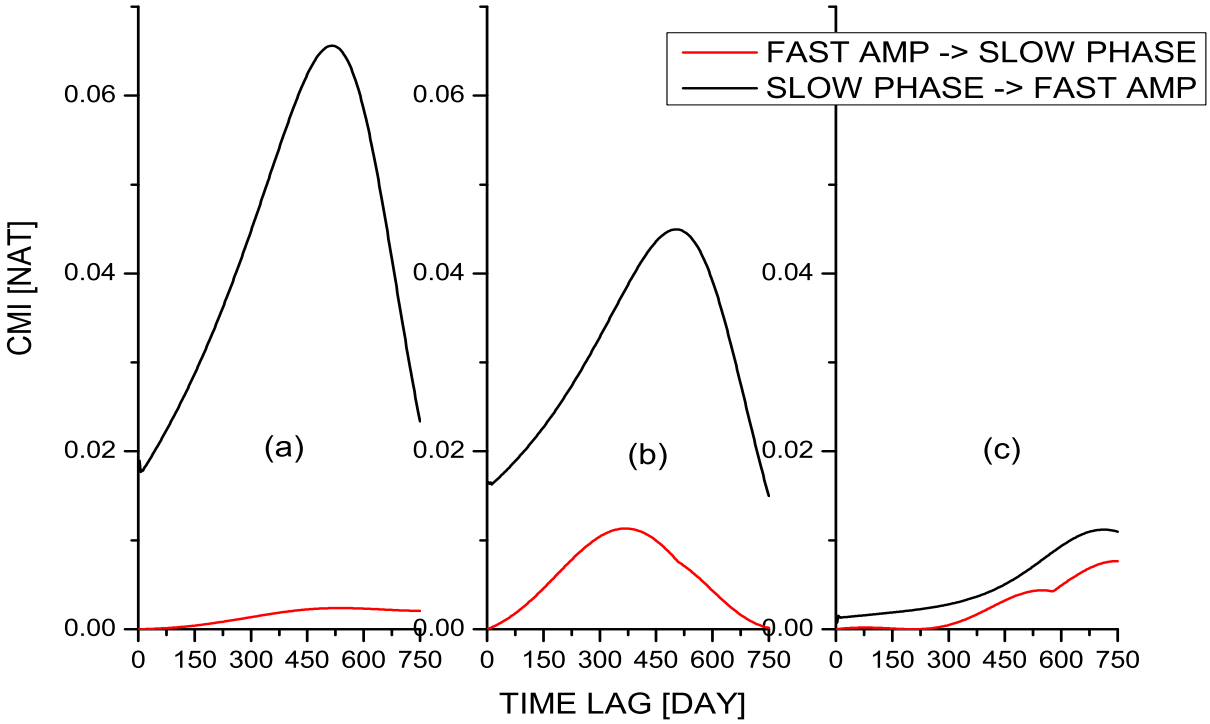

FIG. 2. The conditional mutual information $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), A_2(t-2\eta))$, characterizing the causal influence of the phase $\phi_1(t)$ on the amplitude $A_2(t+\tau)$ (the black curve); and the conditional mutual information $I(A_2(t); \phi_1(t+\tau)|\phi_1(t), \phi_1(t-\eta), \phi_1(t-2\eta))$, characterizing the causal influence in the opposite direction $A_2(t) \to \phi_1(t+\tau)$ (the red curve) as the function of the time lag τ. The periods p_1 for the phase ϕ_1 and p_2 for the amplitude A_2, given as p_1 – p_2 are (a) 8yr – 1.3yr; (b) 6yr – 1.3 yr; and (c) 11yr – 1.3yr. Note that the backward lag η is always set to 1/4 of the period of slow oscillations, i.e., it is 2yr in (a), 1.5yr in (b) and 2.75yr in (c). The SAT data from the Prague-Klementinum station was used.

(1)–(3) in the Letter, one can study various types of cross-scale interactions involving the instantaneous amplitudes, the instantaneous phases, or instantaneous frequencies, obtained as temporal derivatives of the latter. The successful detection of the phase-amplitude interactions, using the estimator (9), means that the amplitudes of the faster variability change within the same time scales as the phase of the slow oscillatory modes, while the high-frequency variability is encoded in the phase of the faster (small time-scale) variability. For detection of all possible interactions one should apply some of general nonlinear estimators which, however, would require even more data for a satisfactory detecting performance [8, 9]. For instance, we have tested a presence of phase-phase interactions using the equiquantal binning algorithm [8], however, no consistent interaction patterns have been observed in the SAT data. It is premature to conclude whether such interactions do not exist, or present methods and available data do not allow a successful detection. Therefore, in the present Letter we report only the phase-amplitude interactions which have been reliably detected.

The time series length equal to the power of two (32768) was chosen because of the application of the discrete Fourier and wavelet transforms, the latter was used for the construction of the wavelet dyadic trees for the multifractal randomization [12].

An example of the statistical evaluation of the CMI $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), A_2(t-2\eta))$, quantifying the causal influence of the phase ϕ_1 on the amplitude A_2 for the phase $\phi_1(t)$ obtained by using the complex continuous wavelet transform with the central period 8 years and the amplitude $A_2(t)$ obtained by using the CCWT with the central period 1.3 years from the Prague-Klementinum daily SAT is presented in Fig. 5. While the CMI value obtained for the original SAT data is 0.0452, the histogram of the CMI values for 1000 realizations of the FT surrogate data is presented in Fig. 5, the left panel. From the cumulative histogram in the right panel of Fig. 5 we can read that the original data CMI value (0.0452) is greater than 99.7% of the surrogate data CMI values. The surrogate data are realizations of the null hypothesis of no cross-scale interactions, i.e. the hypothesis that $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), A_2(t-2\eta)) = 0$.

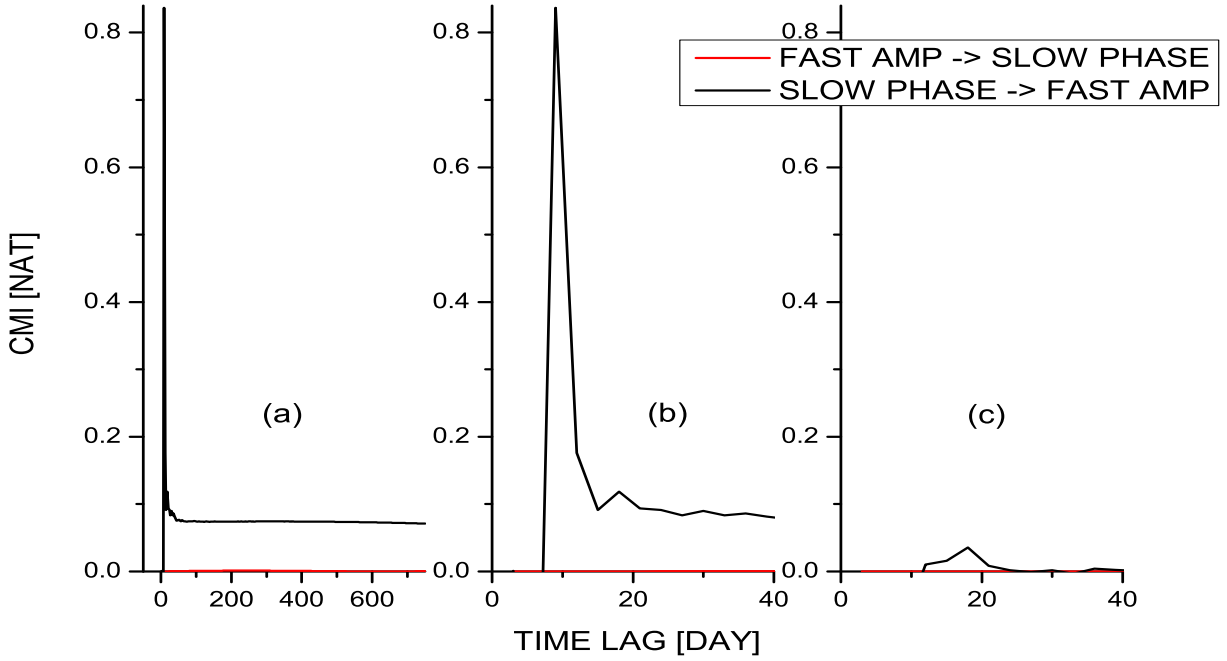

FIG. 3. The same as in Fig. 2, but for the periods (a) 8yr – 2.5yr; (b) 8yr – 2.5 yr (a subset of the time lags); and (c) 6yr – 2.5yr.

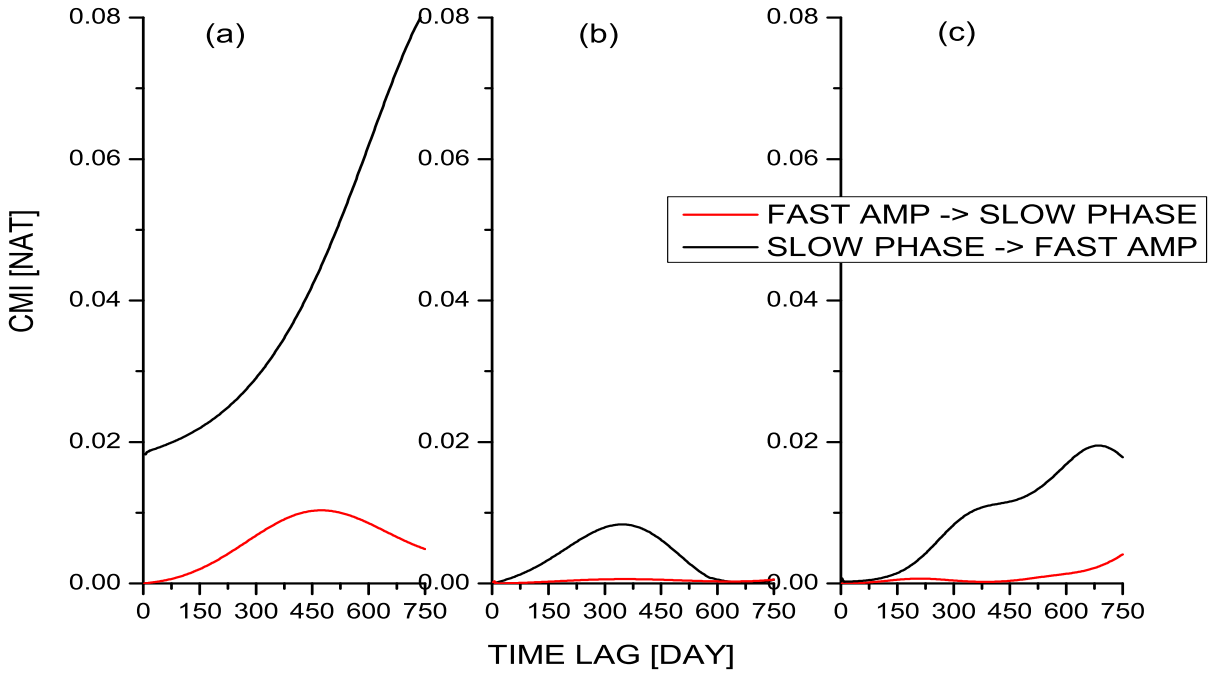

FIG. 4. The same as in Fig. 2, but for the periods (a) 8yr – 0.8yr; (b) 11yr – 0.8 yr; and (c) 7yr – 0.4yr.

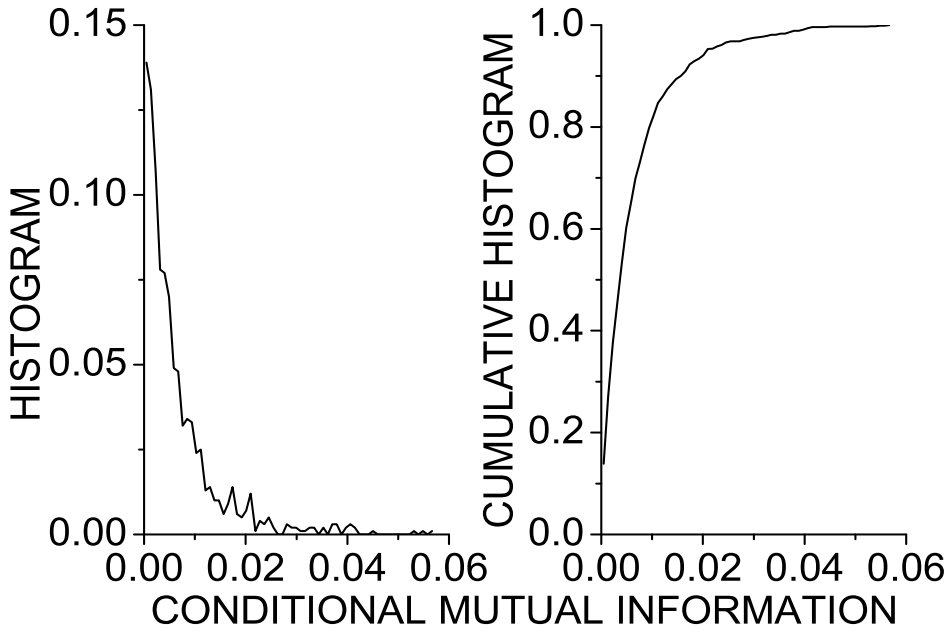

FIG. 5. Histogram (left panel) and cumulative histogram (right panel) of the conditional mutual information $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), A_2(t-2\eta))$ for 1000 realizations of the FT surrogate data. The period for ϕ_1 is 8 years, the period for A_2 is 1.3 years. The corresponding CMI value for the original Prague SAT data is $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), A_2(t-2\eta)) = 0.0452$ which, according to the cumulative histogram, corresponds to the "significance level" 0.997; i.e. the significantly positive CMI establishes the directional $\phi_1 \to A_2$ interaction with $p < 0.003$.

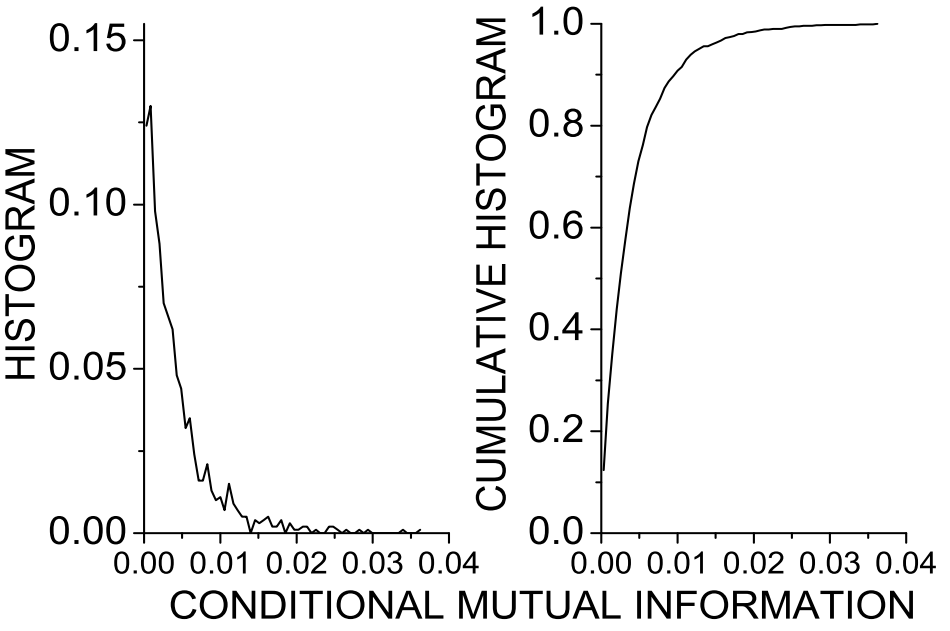

FIG. 6. Histogram (left panel) and cumulative histogram (right panel) of the conditional mutual information $I(A_2(t); \phi_1(t+\tau)|\phi_1(t), \phi_1(t-\eta), \phi_1(t-2\eta))$ for 1000 realizations of the FT surrogate data. The period for ϕ_1 is 8 years, the period for A_2 is 1.3 years. The corresponding CMI value for the original Prague SAT data is $I(A_2(t); \phi_1(t+\tau)|\phi_1(t), \phi_1(t-\eta), \phi_1(t-2\eta)) = 0.00245$ which, according to the cumulative histogram, corresponds to the "significance level" 0.5; i.e. the CMI is not significantly different from zero and no directional $A_2 \to \phi_1$ interaction can be inferred.

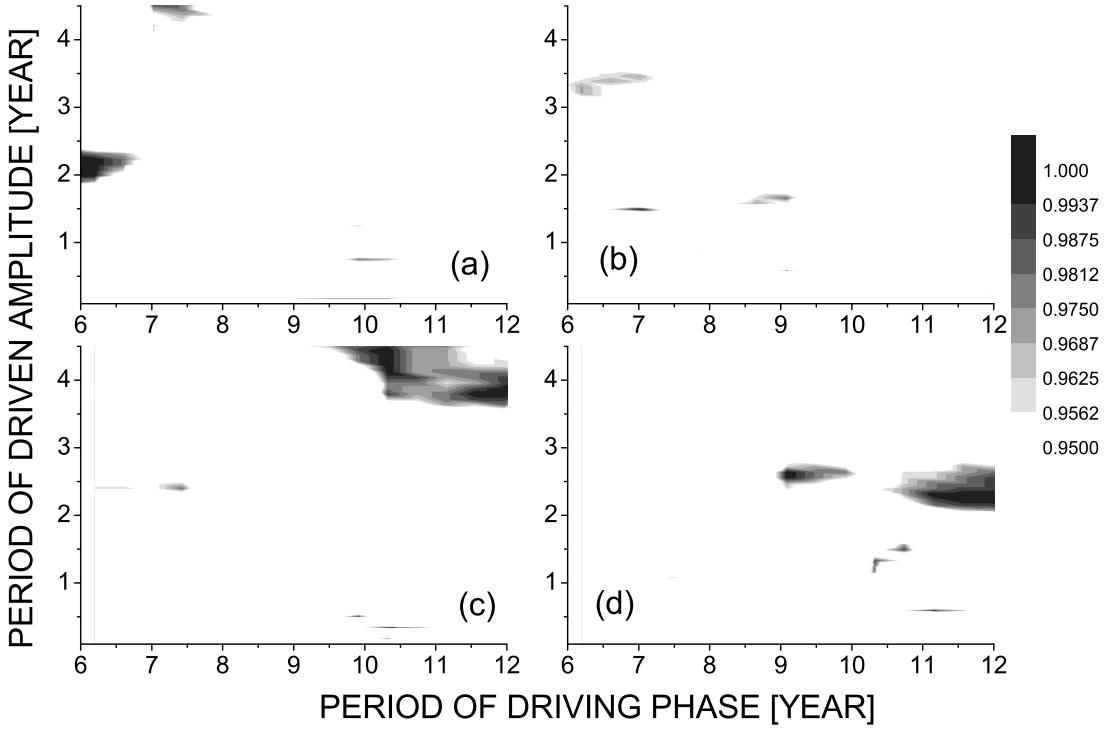

FIG. 7. Causal influence of the phase of slower oscillations on the amplitude of faster fluctuations in different realizations of the MF (a, b) and FT (c, d) surrogate data constructed from the daily surface air temperature from Prague-Klementinum. Here the significance levels greater than 0.95 (grey-coded) demonstrate the false positive results. Note that the CMI is averaged over the forward lags τ 1–750 days, the backward lag η is set to 1/4 of the period of the slower oscillations characterized by the phase ϕ_1.

This null hypothesis is rejected and the data CMI is significantly positive with $p < 0.003$, i.e., there is a probability $p < 0.003$ that the CMI is positive by chance. The "significance level" 0.997 is grey-coded in $\phi_1 \to A_2$ directional interaction charts such as that in Fig. 8.

The statistical evaluation of the interactions in the opposite direction $A_2 \to \phi_1$ reflecting a possible influence of the amplitude A_2 on the phase ϕ_1 for the same pair of periods as above is presented in Fig. 6. For the original Prague SAT data $I(A_2(t); \phi_1(t+\tau)|\phi_1(t), \phi_1(t-\eta), \phi_1(t-2\eta)) = 0.00245$, the significance level is approximately 0.5 and the null hypothesis of no directional $A_2 \to \phi_1$ interaction ($I(A_2(t); \phi_1(t+\tau)|\phi_1(t), \phi_1(t-\eta), \phi_1(t-2\eta)) = 0$) is accepted.

As we can see in Fig. 8 and the related figures of the Letter, we compute the CMI for a range of low and high frequencies. Therefore we encounter the problem of multiple testing. Due to the redundant CCWT decomposition, the tests for subsequent frequencies are not independent. Below we can see that the phases (amplitudes) for any period are correlated with the phases (amplitudes) of a broad range of neighboring periods. At this stage we are not able to determine the number of effectively independent tests, but from the dependence patterns (below) we can guess that it is considerably smaller than the number of the actual tests. Therefore we present the "significance levels" of the single tests and caution that we can encounter false positive results (falsely significant results) at some rate. Plotting the "significance levels" from 0.95, the rate of false positives can reach 5%. In order to study occurrence of such "false positives," in Fig. 7 we present examples of tests in which a realization of surrogate data is used in the role of experimental data, tested by using a set of other surrogate data realizations. The dependent tests due to the CCWT redundancy lead to the observation that the false positives do not occur as randomly distributed points in the plane of the studied ranges of periods, but as randomly placed spots or stains of various shapes and sizes. The FT surrogate data (Fig. 7c, d) apparently suffer from a higher rate of the false positives than the MF surrogate data (Fig. 7a, b).

The significance levels for the CMI $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta))$ obtained from the

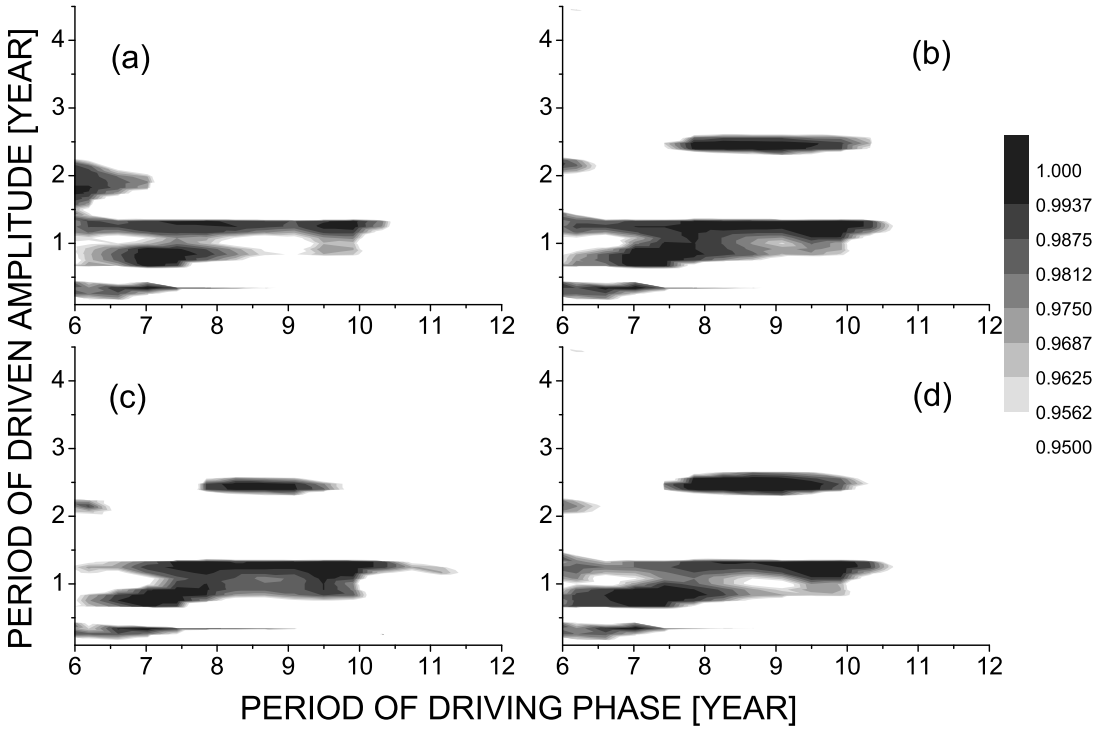

FIG. 8. Causal influence of the phase of slower oscillations on the amplitude of faster fluctuations in the daily surface air temperature from Prague-Klementinum. The significance levels (grey-coded if they are greater than 0.95) for the conditional mutual information $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), \ldots, A_2(t-m\eta))$ with (a) 2-dimensional, (b, d) 3-dimensional, and (c) 4-dimensional condition, obtained using the (a,b,c) Fourier-transform, and (d) multifractal surrogate data. Note that the CMI is averaged over the forward lags τ 1–750 days, the backward lag η is set to 1/4 of the period of the slower oscillations characterized by the phase ϕ_1.

Prague-Klementinum daily SAT data are presented in Fig. 8. The results for the CMI with 2, 3 and 4 conditioning variables are depicted in Fig. 8 a, b, and c, respectively. Since the increase of the conditioning dimensionality over three does not substantially change the patterns of $\phi_1 \to A_2$ interactions, the results presented in the Letter were obtained by evaluating $I(\phi_1(t); A_2(t+\tau)|A_2(t), A_2(t-\eta), A_2(t-2\eta))$ with the 3-dimensional condition. The tests with the two types of the surrogate data bring consistent results (cf. Fig. 8b and Fig. 8d) in spite of the above observation of different rates of false positives. The areas of the significant $\phi_1 \to A_2$ interactions in the SAT data are not only considerably larger than the areas of false positives, but also consistently localized in the SAT data from different locations (cf. the results from Prague in Fig. 8 and from Potsdam and Hamburg in Fig. 1 in the Letter).

No consistent pattern of statistically significant influence has been observed in the opposite direction $A_2 \to \phi_1$ given by $I(A_2(t); \phi_1(t+\tau)|\phi_1(t), \phi_1(t-\eta), \ldots, \phi_1(t-m\eta))$.

IV. FREQUENCY LOCALIZATION

The patterns of statistically significant directional ϕ_1–A_2 interactions (Fig. 8) show localized periods for the driven amplitudes, and a more broad-band structure for the periods of the driving phase. One can ask whether this pattern reflects physical properties of the phenomenon or can be induced by the analytic tools used. The already mentioned redundant character of the used CCWT with an equidistant step in the applied time scales can induce dependence between the instantaneous phases and amplitudes obtained for different time scales (central wavelet periods). In order to quantify this effect, we compute the mutual information between the instantaneous phases (Fig. 9a, c) and between the instantaneous amplitudes (Fig. 9b, d) of the CCWT-extracted oscillatory components with the central wavelet period 8 years (Fig. 9a, b) and 1.3 years (Fig. 9c, d) and the range of periods used in the study of the directional ϕ_1–A_2 interactions (Fig. 8). The black curves illustrate this dependence for the original Prague-Klementinum SAT

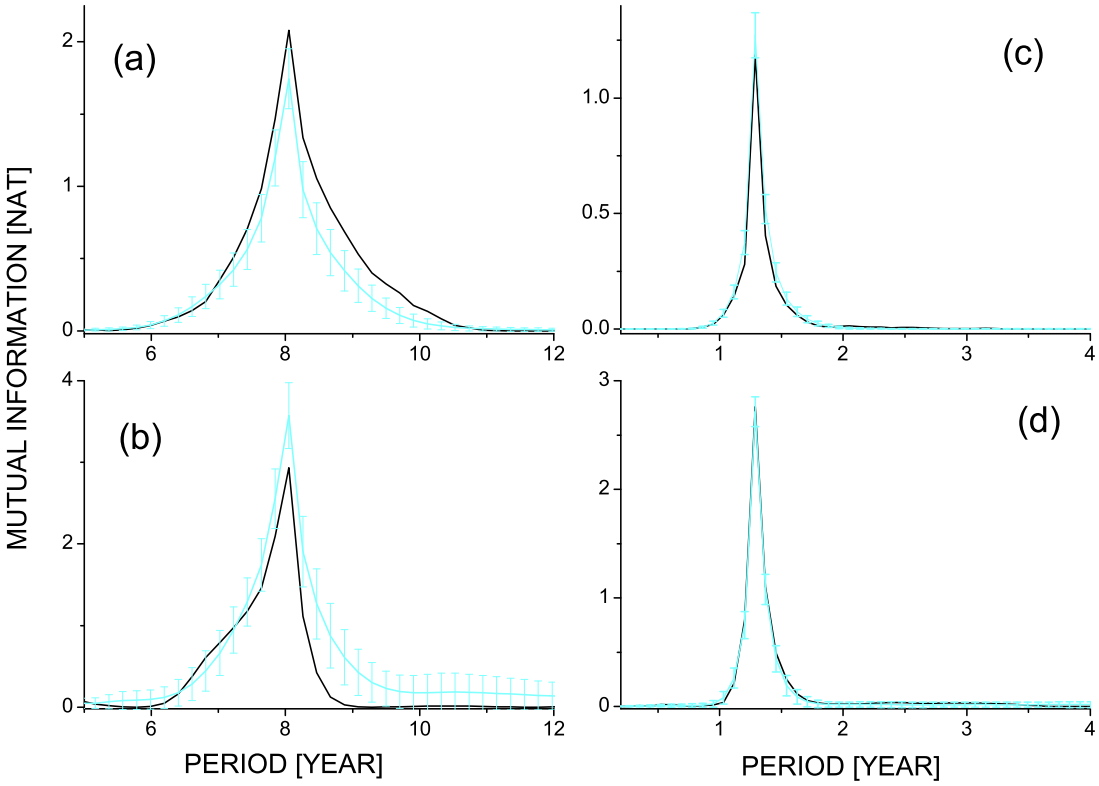

FIG. 9. Dependence (measured by using the mutual information) between instantaneous phases (a, c) and amplitudes (b, d) for the periods on the abscissa and phases/amplitudes for the period 8 years (a, b) and 1.3 years (c, d). The black curves depict the values for the instantaneous phases and amplitudes obtained from the daily surface air temperature from Prague-Klementinum by using the CCWT with the Morlet wavelet. The light blue curves and bars give the mean values and the range of 2 standard deviations, respectively, for 1000 realizations of the FT surrogate data.

data, while the light blue curves depict the mean values for the FT surrogate data with the bars reflecting the range of 2 standard deviations of the surrogate data. The CCWT redundancy apparently induces a certain bandwidth of dependent oscillatory components of close periods.

The bandwidth of the dependent CCWT components is broader when considering the phases than the bandwidth obtained for the amplitudes; and it is broader for larger time scales around the 8-year period than for the smaller time scales around the 1.3-year period. Thus the shape of the ϕ_1–A_2 interaction pattern can be artificially extended over a range of the slower driving periods. On the other hand, we can compare the phase/amplitude dependence results for the original SAT data and their FT surrogate data. While there are no differences for smaller time scales around the 1.3-year period (Fig. 9c, d), for the larger time scales around the 8-year period the results from the SAT data differ from the FT surrogates even with a statistical significance (considering the 2SD range of the surrogates, Fig. 9a, b). The phase coherence over the neighboring periods is broader in the SAT data than in the surrogates, while the amplitude dependence in the SAT data is narrower than in the FT surrogate data. This behavior reminds the property of phase-synchronized systems with coherent (dependent) phases but independent amplitudes [14]. The statistically significant differences in Fig. 9a also constitute a new test for nonlinearity for the oscillatory phenomenon acting on the periods around 8 years. This phenomenon induces coherent phases to oscillatory components in a broader range than the range of phase coherence induced by the used CCWT decomposition applied to realizations of a linear stochastic process with the same frequency spectrum as the original SAT data. This analysis supports the conclusion that we observe a nonlinear, possibly broadband phenomenon, acting in a range of periods between 6 and 11 years and influencing temperature variability

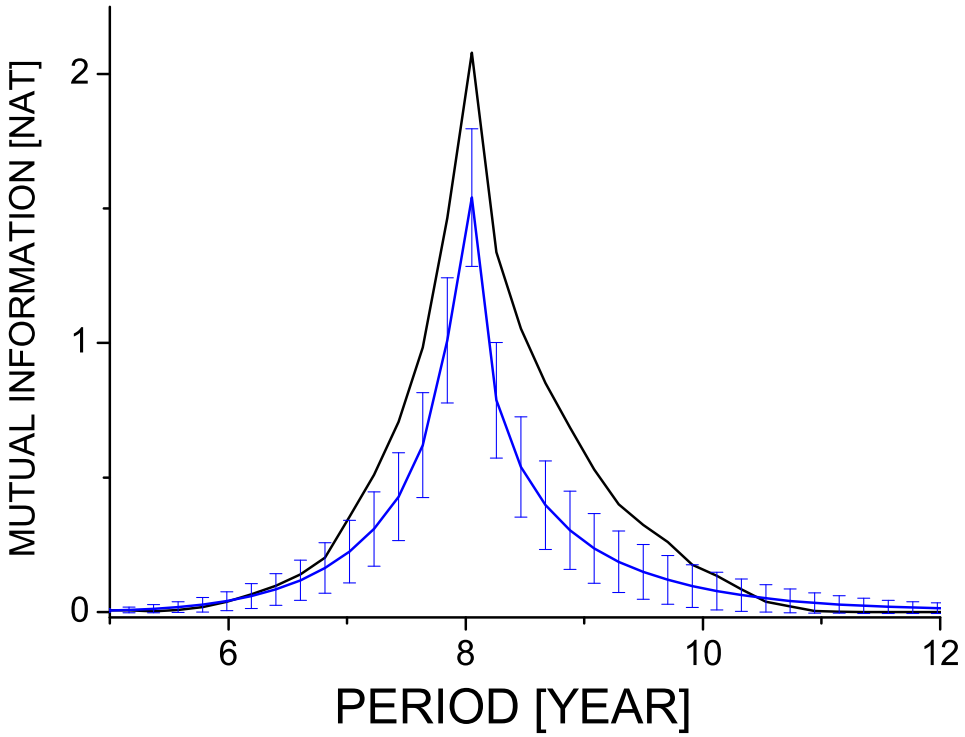

FIG. 10. Dependence (measured by using the mutual information) between the instantaneous phases for the periods on the abscissa and the instantaneous phases for the period 8 years. The black curve depicts the values for the instantaneous phases obtained from the daily surface air temperature from Prague-Klementinum by using the CCWT with the Morlet wavelet. The blue curve and bars give the mean values and the range of 2 standard deviations, respectively, for 1000 realizations of independent, identically, normally distributed 32768-sample data (Gaussian white noise).

on shorter time scales. We should not forget, however, the possibility that the used CCWT decomposition might artificially extend the bandwidth of phase-coherent oscillations, since the FT surrogates should produce a set of independent oscillatory components. In order to see the "net" effect of the limited frequency localization due to CCWT without possible confounding effects of the FT surrogate algorithm, in Fig. 10 using the blue curves and bars we illustrate the range of false phase coherence due to the CCWT decomposition applied to white Gaussian noise. Thus the actual multiscale phenomenon in the atmospheric dynamics can have a narrower bandwidth than the observed range of periods 6–11 years, or, two or more distinct phenomena can be merged together due to a limited frequency localization of the used CCWT decomposition. The CCWT frequency localization can be increased by tuning the mother wavelet parameters. However, according to the uncertainty principle, sharp localization in time and frequency are mutually exclusive [15, 16]. Increasing the localization in frequency we would decrease the localization in time which would effect the test for causal ϕ_1–A_2 interactions. The problem of increasing the frequency localization without loosing the sensitivity of the test for the directional ϕ_1–A_2 interactions will be one of the most important topics of further research.

V. NORTH ATLANTIC OSCILLATION AS A POSSIBLE CLUE FOR A MECHANISM OF CROSS-SCALE INTERACTIONS

Oscillatory phenomena with the periods between 6 and 11 years, however, most frequently with the period around 7–8 years have been observed in the air temperature and other meteorological data such as sea surface temperature and salinity, Baltic Sea ice annual maximum extent, precipitation records, river runoff records, and most importantly, in the index of the North Atlantic Oscillation (see Ref. [17, 18] and references therein).

The North Atlantic Oscillation (NAO) is a dominant pattern of atmospheric circulation variability in the extratropical Northern Hemisphere and is a major factor influencing meteorological variables including tempera-

ture, precipitation, occurrence of storms, wind strength and direction in the Atlantic sector and surrounding continents. On the global scale, NAO has a climate significance that rivals the Pacific El Niño Southern Oscillation (ENSO) [19]. Since NAO reflects synchronous variations of the pressure gradient between the Icelandic Low and the Azores High on timescales from daily to multidecadal, the NAO phenomenon is characterized by the NAO index which is defined as the normalized pressure difference between the Azores and Iceland. The character of the NAO index time series is dominated by a colored-noise process with a 1/f spectrum, however, subtle analysis methods such as the Monte Carlo Singular Spectrum Analysis (see [17] and references therein) can uncover quasi-oscillatory phenomena explaining a part of its variability. Gámiz-Fortis et al. [20] observed a cycle with a period 7.7 years in the winter NAO index. Paluš & Novotná [17] have also detected the 7–8-year cycle in the (whole-year) NAO index and found that it is phase-synchronized with the equivalent cycle in SAT recordings from a number of European locations. There are indications that the 7–8 year cycle in the North Atlantic has its roots in nonlinear ocean-atmosphere interactions. According to Feliks et al. [21, 22] the 7-8 year cycle might be induced by an oscillation of a similar period in the position and strength of the Gulf Stream's sea surface temperature front in the North Atlantic. The 7-8 year variability in the Gulf Stream, in turn, has been attributed to an oscillatory gyre mode of the North Atlantic's wind-driven circulation.

In the present Letter we have studied the atmospheric cross-scale interactions as local phenomena in SAT recordings from individual locations. The roots of these phenomena, however, might stem from large-scale interactions in which dominant modes of circulation variability, such as the NAO, play important roles. A new perspective in research of climate variability has been opened by using the tools from complex network theory in analysis of multivariate climate data such as the gridded SAT reanalysis [23]. Changes in the NAO induce changes in connectivity of climate networks [24, 25]. When the NAO index is distinctly positive (so-called positive NAO phase, NAO+) we can observe an increased network connectivity in large areas of the Northern Hemisphere, including the European areas presented in Fig. 4 of the Letter. Effectively it means that correlations between temperature records from different locations are stronger during the NAO+ periods, that during the NAO– periods (defined in full analogy). Hurrell and Dickson [26] help us to understand this observation: "in the so-called positive phase, higher than normal surface pressures south of 55°N combine with a broad region of anomalously low pressure throughout the Arctic and subarctic. Consequently, this phase of the oscillation (NAO+) is associated with stronger-than-average westerly winds across the middle latitudes of the Atlantic onto Europe, with anomalous southerly flow over the eastern United States and anomalous northerly flow across western Greenland, the Canadian Arctic, and the Mediterranean. The easterly trade winds over the subtropical North Atlantic are also enhanced during the positive phase of the oscillation. During the negative phase (NAO–), both the Icelandic low- and Azores high-pressure centers are weaker-than-normal, so both the middle latitude westerlies and the subtropical trade winds are also weak."

The NAO is a dominant mode of low-frequency atmospheric variability [27]. Its changes on larger time scales influence the character of atmospheric circulation and thus the atmospheric dynamics on shorter time scales. Considering the observation that the low-frequency dynamics in the NAO and in the air temperature in large areas of Europe are phase-synchronized [18], we can suspect that the NAO and its large-scale interactions are important for understanding of mechanisms of the observed local cross-scale interactions. These considerations, however, are very preliminary, and the cross-scale interactions in the multiscale atmospheric dynamics require further research in all aspects of theory, modelling and data analysis.

VI. ACKNOWLEDGEMENTS

The author would like to thank the anonymous referees who have provided insightful comments and useful suggestions which helped to improve this paper.

This study was supported by the Czech Science Foundation project No. P103/11/J068.

[1] M. Paluš, Phys. Lett. A **213** 138 (1996).
[2] T.M. Cover and J.A. Thomas, *Elements of Information Theory* (J. Wiley & Sons, New York, 1991).
[3] M. Paluš, V. Albrecht, I. Dvořák, Phys. Lett. A **175** 203 (1993).
[4] M. Paluš, V. Komárek, Z. Hrnčíř and K. Štěrbová, Phys. Rev. E **63** 046211 (2001).
[5] F. Takens, In: D.A. Rand and D.S. Young (Editors), *Dynamical Systems and Turbulence (Warwick 1980)*, Lecture Notes in Mathematics 898. (Springer, Berlin, 1981), pp. 366–381.
[6] T. Schreiber, Phys. Rev. Lett. **85** 461 (2000).
[7] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka and J. Bhattacharya, Phys. Rep. **441** 1 (2007).
[8] M. Paluš, and M. Vejmelka, Phys. Rev. E **75** 056211 (2007). doi: 10.1103/PhysRevE.75.056211
[9] M. Vejmelka, and M. Paluš, Phys. Rev. E **77** 026214 (2008). doi: 10.1103/PhysRevE.77.026214
[10] D. Prichard, and J. Theiler, Physica D **84** 476-493 (1995). doi: 10.1016/0167-2789(95)00041-2

[11] A. Molini, G.G. Katul, and A. Porporato, J. Geophys. Res. **115** D14123 (2010). doi:10.1029/2009JD013016.
[12] M. Paluš, Phys. Rev. Lett. **101** 134101 (2008). doi: 10.1103/PhysRevLett.101.134101
[13] C. Torrence, and G.P. Compo, Bull. Amer. Meteorological Society **79** 61 (1998).
[14] M.G. Rosenblum, A.S. Pikovsky, J. Kurths, Phys. Rev. Lett. **76,** 1804 (1996).
[15] M. Vetterli, J. Kovacevic, Wavelets and Subband Coding, Prentice Hall, Englewood Cliffs, 1995.
[16] S. Mallat, A Wavelet Tour of Signal Processing, 3rd edn., Academic Press, Burlington, 2008.
[17] M. Paluš, and D. Novotná, J. Atmos. Sol.-Terr. Phys. **71** 923 (2009). doi:10.1016/j.jastp.2009.03.012.;
[18] M. Paluš, and D. Novotná, Nonlin. Processes Geophys. **18** 1 (2011) doi:10.5194/npg-18-1-2011.
[19] J. Marshall, Y. Kushnir, D. Battisti, et al., Int. J. Clim., **21** 1863–1898 (2001) doi:10.1002/joc.693.
[20] S.R. Gámiz-Fortis, D. Pozo-Vázques,M.J. Esteban-Parra, and Y. Castro-Díez, J. Geophys. Res. **107(D23)** 4685 (2002).
[21] Y. Feliks, M. Ghil, and A.W. Robertson, J. Clim. **23** 4060–4079 (2010). doi:10.1175/2010JCLI3181.1.
[22] Y. Feliks, M. Ghil, and A.W. Robertson, J. Clim. **24** 522–542 (2011). doi: 10.1175/2010JCLI3859.1.
[23] A. Tsonis, and P. Roebber, Physica A **333** 497–504 (2004). doi: 10.1016/j.physa.2003.10.045.
[24] M. Paluš, D. Hartman, J. Hlinka, and M Vejmelka, Nonlin. Processes Geophys. **18** 751–763 (2011). doi:10.5194/npg-18-751-2011.
[25] O. Guez, A. Gozolchiani, Y. Berezin, S. Brenner, and S. Havlin, EPL **98** 38006 (2012). doi: 10.1209/0295-5075/98/38006
[26] J. Hurrell, and R. Dickson, Climate variability over the North Atlantic, in: Marine ecosystems and climate variation: the North Atlantic: a comparative perspective, edited by Stenseth, N., Ottersen, G., and Hurrell, J., pp. 15–31, Oxford University Press, USA, 2005.
[27] J. W. Hurrell, Y. Kushnir, and M. Visbeck, Science **291(5504)** 603–605 (2002).