

Estimating causal interaction between prefrontal cortex and striatum by transfer entropy

Chaofei Ma · Xiaochuan Pan · Rubin Wang ·
Masamichi Sakagami

Received: 16 October 2012/Revised: 21 December 2012/Accepted: 25 December 2012/Published online: 4 January 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Transfer entropy (TE) is an information-theoretic measure for the investigation of causal interaction between two systems without a requirement of pre-specific interaction model (such as: linear or nonlinear). We introduced an efficient algorithm to calculate TE values between two systems based on observed time signals. By this method, we demonstrated that the TE correctly estimated the coupling strength and the direction of information transmission of two nonlinearly coupled systems. We also calculated TE values of real local field potentials (LFPs) recorded simultaneously in the lateral prefrontal cortex (LPFC) and the striatum of the behavioral monkey, and observed that the TE value from the LPFC to the striatum was stronger than that from the striatum to the LPFC, consistent with anatomical structure between the two areas. Moreover, the TE value dynamically varied dependent on behavior stages of the monkey. These results from simulated and real LFPs data suggested that the TE was able to effectively estimate functional connectivity between different brain regions and characterized their dynamical properties.

Keywords Causal interaction · Transfer entropy · Mutual information · Local field potential

Introduction

The brain is a complex system consisting of interconnected modules (such as: cortical areas) that often perform very specific operation. For instance, primary sensory areas receive sensory information and transfer it to higher associated areas. Signals from higher-order areas are feedback to primary sensory areas to influence their information processing, forming brain networks with forward and feedback connections. A fundamental problem in cognitive neuroscience is to clarify how information is processed in such networks and how the networks generate corresponded cognitive functions. Many researchers record neural signals simultaneously from various brain regions to investigate these issues. For example, in animal's neurophysiological experiments, neural spikes and local field potentials (LFPs) are able to be recorded at the same time in multiple brain areas using multi-electrode arrays. Also by invasive methods, we can record neural signals from sculpt, such as electroencephalography (EEG), magnetoencephalography (MEG) (Gu and Liang 2007). Based on these recorded data, one question is how to measure and estimate the strength of functional connectivity and information flow, or causal interaction, between brain regions (Fingelkurts et al. 2005). One popular way of augmenting the concept of causality was introduced by Norbert Wiener (Wiener 1956). In Wiener's definition a variable X_1 causes a variable X_2 if information in the past of X_1 helps predict the future of X_2 with better accuracy than only information in the past of X_2 itself.

So far most implementation of Wiener's principle used model-based approaches, such as Granger-causality

C. Ma
Institute for Cognitive Neurodynamics,
School of Information Science and Engineering, East China
University of Science and Technology, Shanghai 200237,
People's Republic of China

X. Pan (✉) · R. Wang
Institute for Cognitive Neurodynamics, School of Science,
East China University of Science and Technology,
200237 Shanghai, People's Republic of China
e-mail: pxc@ecust.edu.cn

M. Sakagami
Brain Science Institute, Tamagawa University, Tokyo, Japan

(Kaminski et al. 2001; Freeman 2007; Seth and Edelman 2007; Seth 2008; Barnett et al. 2009; Werner 2009; Hu et al. 2012), to analyze causal interaction in brain networks. However, Granger-causality and related methods come with a serious drawback as they specify a linear model of interaction a priori (Granger 1969; Ding et al. 2000). It is common that brain networks have nonlinear behavior of its parts and nonlinear interactions between them. To characterize nonlinear interactions between two systems, many researchers have used the concept of mutual information (Tsukada et al. 1975, 1976; Dan et al. 1998; Paninski 2003), which is quantified in Shannon's information theory (Shannon and Weaver 1949). Unfortunately, mutual information does not provide any dynamical and directional information. Therefore, Schreiber proposed a method termed transfer entropy (TE) on the basis of Wiener's principle of causality to estimate causal interaction between two systems (Schreiber 2000). TE is an information theoretical measure able to estimate causal relationship from time series. It takes into account linear and nonlinear interactions, without the use of *a priori* specification of the interaction mechanism itself (Schreiber 2000). Thus TE can represent a general way to define the causal strength between brain regions.

TE has been widely used to analyze neural data (Gourévitch and Eggermont 2007; Garofalo et al. 2009; Besserve et al. 2010; Ito et al. 2011; Lindner et al. 2011). But to date most of its applications concentrate spike trains that consist of a sequence of two states of "1" and "0" (Gourévitch and Eggermont 2007; Garofalo et al. 2009; Ito et al. 2011). It is usually easy to estimate the state probability and the transition probability in the spike trains that are necessary to calculate the TE values. So far, there are few studies that apply the method of TE to analyze continuous neural data such as LFP, EEG and MEG (Besserve et al. 2010; Lindner et al. 2011). One difficult question is how to calculate the state and transition probabilities from time series.

We used an efficient approach to compute TE in this paper. We first embedded raw data of time series into a state space with high dimension, and then applied nearest-neighbor techniques to estimate the state probability and the transition probability. We utilized this method to calculate TE values of two simulated time series with nonlinear coupling, and found that the TE value increased with the increment of the coupling strength between the two signals, and decreased with the increasing strength of noise. We also calculated the TE value of LFPs recorded simultaneously from the lateral prefrontal cortex (LPFC) and the striatum of a monkey performing a stimulus-stimulus association task. The results showed that the TE value from the LPFC to the striatum was larger than the TE value from the striatum to the LPFC, consistent with anatomical evidence that prefrontal neurons have direct

projections to striatal neurons, but striatal neurons do not (Alexander et al. 1986). Our results demonstrated that the TE was able to correctly estimate the strength of functional connectivity and the direction of information flow between areas in the brain network.

Modeling

Mutual information

Definition: Assuming there are two random variables (I, J), their joint probability distribution is $p(i, j)$, and marginal probabilities are $p(i), p(j)$, respectively. The mutual information $M(i, j)$ is defined as relative entropy of the joint distribution $p(i, j)$ and the product of the marginal distributions $p(i) p(j)$ (Shannon and Weaver 1949), namely:

$$M_{IJ} = \sum p(i, j) \log \frac{p(i, j)}{p(i) \cdot p(j)} \quad (1)$$

Transfer entropy

There are two stochastic time series I, J , their joint probability distribution is $p(i, j)$, and the conditional probability is $p(i|j)$. The conditional entropy can be defined as the following equation (Shannon and Weaver 1949):

$$H_{I|J} = - \sum p(i, j) \log p(i|j) = H_{IJ} - H_J \quad (2)$$

From Eq. (2), we can get a relation: $H_{I|J} - H_{J|I} = H_I - H_J$, so the conditional entropy is asymmetrical. This asymmetry, however, is only due to the difference of individual Shannon entropies and not due to information flow. Mutual information could have directional information by introducing a time lag τ in either one of the variables, as the following equation,

$$M_{IJ} = \sum p(i_n, j_{n-\tau}) \log \frac{p(i_n, j_{n-\tau})}{p(i) \cdot p(j)} \quad (3)$$

Now we consider a dynamic system by studying its transition probabilities rather than static probabilities. We assume that two time series $I = i_n$ and $J = j_n$ can be approximated by Markov processes, and measure the deviation from the following generalized Markov process:

$$P(i_{n+1}|i_n^k, j_n^l) = P(i_{n+1}|i_n^k) \quad (4)$$

where $i_n^k = (i_n, \dots, i_{n-k+1})$ and $j_n^l = (j_n, \dots, j_{n-l+1})$, while k and l are the orders of Markov processes of I and J . When the transition probabilities or dynamics of I are independent of the past of J , the Eq. (4) is fully satisfied, indicating there is no directed interaction from J to I . The TE is defined as Kullback–Leibler divergence (1951) between the two probability distributions at each side of the Eq. (4), as following equation (Schreiber 2000)

$$T_{J \rightarrow I} = \sum p(i_{n+1}, i_n^k, j_n^l) \log \frac{p(i_{n+1} | i_n^k, j_n^l)}{p(i_{n+1} | i_n^k)} \quad (5)$$

where $p(i_{n+1}, i_n^k, j_n^l)$ is transition probability of from state (i_n^k, j_n^l) to state i_{n+1} , $p(i_{n+1} | i_n^k, j_n^l)$ and $p(i_{n+1} | i_n^k)$ are conditional probabilities. Noting that Eq. (5) implies a prediction time of 1 (from i^n to i^{n+1}). In a general way, there may be a longer time delay for interaction between state $i_n^k = (i_n, \dots, i_{n-k+1})$ and state $j_n^l = (j_n, \dots, j_{n-l+1})$. The TE including prediction time u can be defined as:

$$T_{J \rightarrow I} = \sum p(i_{n+u}, i_n^k, j_n^l) \log \frac{p(i_{n+u} | i_n^k, j_n^l)}{p(i_{n+u} | i_n^k)} \quad (6)$$

$T_{J \rightarrow I}$ is now asymmetric, which is determined by the nature of the conditional entropy. The TE also contains directional and dynamic information. It measures the degree of dependence of I on J and not vice versa, and is calculated on the basis of the transition probability.

In order to compute the TE, we first have to estimate the transition probability. Schreiber (2000) suggested that the transition probability could be calculated based on the joint probability and kernel estimation. However, the probability density in stochastic systems is uncertain, and it is difficult to be estimated. The method to calculate the joint probability and kernel estimation is dependent on each individual system.

To calculate the probability density, one efficient way is to reconstruct the state space of the raw data, embedding the scalar time series into trajectories in a state space of possibly high dimension (Lindner et al. 2011). The mapping uses delay-coordinates to create a set of vectors or points in a higher dimensional space according to:

$$X_t^d = (X(t), X(t - \tau), X(t - 2\tau), \dots, X(t - (d - 1)\tau)) \quad (7)$$

where d is the dimension of high-dimensional space and τ denotes delay time. The mapping procedure depends on these two parameters.

After having reconstructed the state spaces of any pair of time series, we estimated the TE between their underlying systems. We rewrite Eq. (6) as sum of four Shannon entropies as Eq. (8):

$$T(X \rightarrow Y) = S(y_t^{d_y}, x_t^{d_x}) - S(\gamma_{t+u}, y_t^{d_y}, x_t^{d_x}) + S(\gamma_{t+u}, y_t^{d_y}) - S(y_t^{d_y}) \quad (8)$$

From Eq. (8), the key problem in calculating the TE is to compute the combination of different joint and marginal differential entropies. We used the nearest-neighbor method and Kraskov-Stögbauer-Grassberger estimator to estimate these probabilities (Duda et al. 2004; Kraskov 2004; Kraskov et al. 2004). After taking into account these methods, we rewrote the TE formula as the following equation:

$$T(X \rightarrow Y) = \Psi(k) + \langle \Psi(n_{y_t^{d_y}} + 1) - \Psi(n_{\gamma_{t+u}, y_t^{d_y}} + 1) - \Psi(n_{y_t^{d_y}, x_t^{d_x}}) \rangle t \quad (9)$$

where Ψ denotes the digamma function, the angle brackets indicate an averaging over different time points. The distances to the k -th nearest neighbor in the highest dimensional space (spanned by $y_{t+u}, y_t^{d_y}, x_t^{d_x}$) define the radius of the spheres for the counting of points n_Z in all the marginal spaces Z involved.

Cao criterion

Taking together, in order to calculate the TE, we first have to estimate d and τ from raw time series. In this paper, we used Cao (1997) criterion to determine the minimum embedding dimension and the delay time. A variation reflecting a relative increase in distance between two nearest neighbors from d to $d + 1$ dimension is defined as:

$$a(t, d) = \left\| \frac{x_t^{d+1} - x_{t'(t,d)}^{d+1}}{x_t^d - x_{t'(t,d)}^d} \right\| \quad (10)$$

where $t = 1, 2, \dots, N - d\tau$, and $\|\cdot\|$ indicates Euclidean distance in d and $d + 1$ dimensions. The vector x_t^d and $x_{t'(t,d)}^d$ are nearest neighbors in the d -dimensional space.

A quantity $E(d)$ averaged across all N instances of $a(t, d)$ is utilized to define the minimum embedding dimension:

$$E(d) = \frac{1}{N - d\tau} \sum_{t=1}^{N-d\tau} a(t, d) \quad (11)$$

$E(d)$ depends only on the dimension d and the time lag τ . The difference of $E(d)$ from d to $d + 1$ dimension is defined as:

$$E1(d) = \frac{E(d+1)}{E(d)} \quad (12)$$

We construct a target function as following:

$$E1(d - 1) + E1(d + 1) - 2 * E1(d) \quad (13)$$

By minimizing this target function, the optimal embedding dimension d and the time lag τ are able to be determined.

Results based on simulated data

To investigate the effectiveness of the proposed algorithm, we first calculated TE values of two simulated time series with nonlinear coupling. The two time series X and Y were constructed through the following two equations:

$$\begin{aligned}
 X(t+1) &= \sum_{i=0}^9 \alpha_i X(t-i) + \varepsilon \eta_x(t) \\
 Y(t+1) &= Y(t) \gamma X(t+1-\delta) + \sum_{i=0}^9 \beta_i Y(t-i) + \varepsilon \eta_y(t)
 \end{aligned}
 \tag{14}$$

where η_x and η_y both are white noise processes; γ is the coupling coefficient of the two time series; δ is the coupling delay; α_i and β_i are the combination coefficients ($i = 0, \dots, 9$); ε is the noise intensity. In Eq. (14), the sequence X contained only its own past information, while the sequence Y contained not only its own past information, but also past information from the sequence X . In another words, there was transferred information from X to Y , but no information from Y to X .

We examined how the parameters, such as the prediction time u , the noise intensity and the coupling strength affected TE values of the two time series. In the simulation, we constructed 100 data sets, and each data set contained 40 trials, and each trial had 1,000 sample points.

We first examined influence of the prediction time on the TE. The simulated data was generated with following parameters: the coupling strength was 1.0, the noise intensity was 0.1, and the delay time was 20 ms. The optimal dimension of embedded space was $d = 4$ and the optimal delay $\tau = 1$ using the Cao criterion. We varied prediction time u from 10 to 32, as $u = \{10, 11, 12, \dots, 32\}$. In the simulation, we needed to verify whether the TE calculated in each data set was significantly different from the null hypothesis that there was no interaction between those data. We used the permutation test. The permutation test is a non-parametrical statistical significance test. The basic approach is following: shuffle 40 trials data in each dataset, split the data into two groups, calculate the TE based on reshuffled data, and repeat this process for 100 times. This gives the distribution of the test statistic under the null hypothesis of no interaction. If the TE value calculated from the original dataset significantly differs from the null distribution ($P < 0.05$), we can reject the null hypothesis, indicating that there is transmission of information between the two time series. We calculated the TE value for each of the 100 data sets and checked whether it was statistically significant or not. The result was shown in Fig. 1.

In Fig. 1, the x-axis represents the prediction time; the y-axis represents the percentage of significant TE values. It can be seen that when u was from 15 to 23, the percent of significant TE values from the sequence X to the sequence Y reached one hundred, while the percentage of significant TE values from the sequence Y to the sequence X was very low. This was because in the simulation data, there was information transmission from the sequence X to the

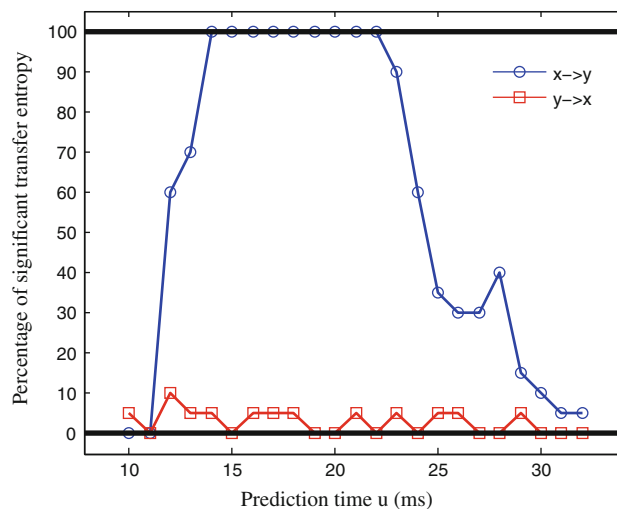


Fig. 1 The percentage of the significant TE value as a function of the prediction time u . The blue curve shows the significant percentage of TE value from the sequence X to the sequence Y , and the red curve indicates the significant percentage of TE value from the sequence Y to the sequence X . (Color figure online)

sequence Y , but had no information transmission from the sequence Y to the sequence X . And the delay time was 20 ms, falling in the range of prediction time that showed one hundred percentage significant TE values. Figure 2 presents the TE strength between the two series as a function of the prediction time u .

In Fig. 2, the x-axis represents the prediction time u ; the y-axis is the TE strength. The TE strength from the sequence X to the sequence Y reached the peak when the prediction time was about 21. While the TE strength from sequence Y to sequence X was around zero. We knew there were 20 ms delays in information transmission from the

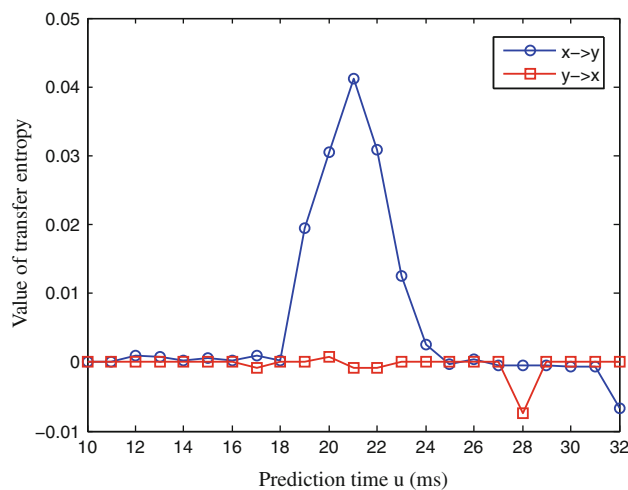


Fig. 2 The TE value as the function of the prediction time u . The blue curve represents the TE value of the sequence X to the sequence Y and the red curve is the TE value of the sequence Y to the sequence X . (Color figure online)

sequence X to the sequence Y . The delay time fell in the effective range of the prediction time, indicating that the TE could estimate the propagation delay between two signals.

Neural signals always contain noise, so we examined how noise affected the TE. We generated the data sets with various strength of noise, and calculated the TE value of each data set. The relation between the strength of noise and the significant percentage was shown in Fig. 3.

With the increase of the noise strength from 0.1 to 0.7, the permutation test detected 100 percentage significant TE values from the series X to the series Y . With further increasing, the significant percentage decreased. In the case of from the sequence Y to the sequence X , we detected a very low percentage of significant TE values. Figure 4 shows the function of TE values against the strength of noise. In the case of from the series X to the series Y the TE value decreased gradually with increasing the strength of noise. In the case of from the series Y to the series X , the TE value was unchanged, around zero. These results demonstrated that the TE was robust to a certain range of noise, indicating the TE method can also be effective in detecting the transmission of information between the two signals even with noise.

We calculated both TE (Fig. 5a) and mutual information (Fig. 5b) of these simulated data. Figure 5a shows the relation between the TE values with the coupling strength. The abscissa in this figure indicates the coupling strength between the two signals; the ordinate represents the TE value. The TE value from the sequence X to the sequence Y increased monotonically with the increasing of coupling

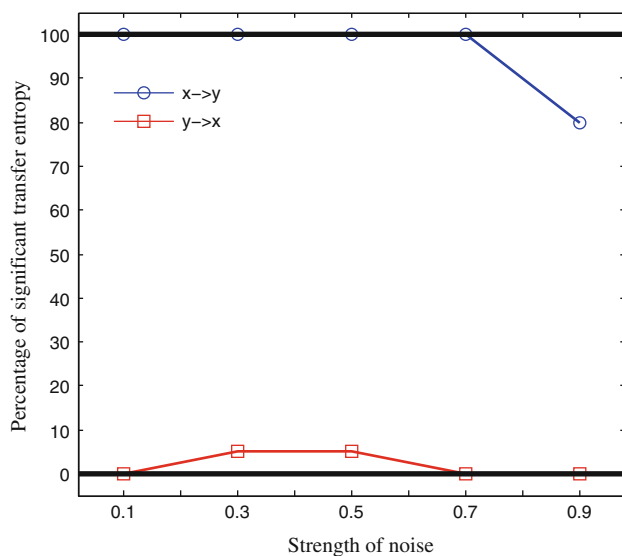


Fig. 3 The effect of the noise intensity on the percentage of significant TE values. The blue curve shows the TE value from the sequence X to the sequence Y , and the red curve represents the TE values from the sequence Y to the sequence X . (Color figure online)

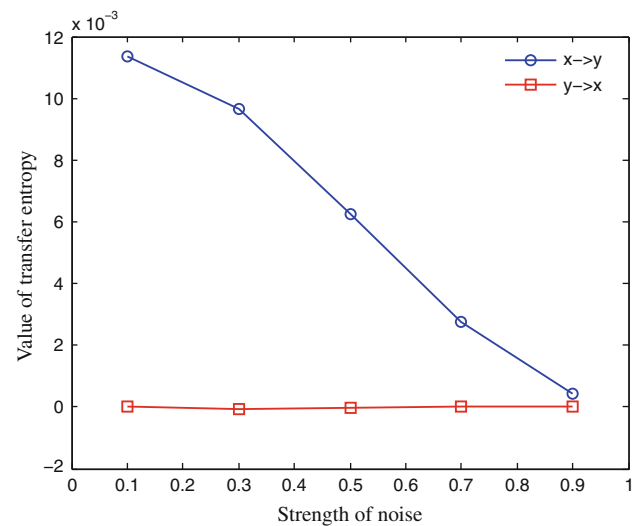


Fig. 4 The effect of the noise intensity on the TE values. The blue curves indicates the TE value from the sequence X to the sequence Y , and the red curves the TE value from the sequence Y to the sequence X . (Color figure online)

strength. On the other hand, the TE value from the series Y to the series X was not different from zero because there was no information was transferred from Y to X . Figure 5b shows mutual information as the function of the couple strength. Both mutual information from X to Y and mutual information from Y to X increased with the coupling strength between the two signals. However, the curves of mutual information in the two directions were overlapped, even though the information transmission between the two signals was asymmetric. Mutual information can detect information transmission between series X and series Y , but not the direction of information flow. Based on the simulated data, we demonstrated that the TE method was able to correctly estimate the coupling delay time, the direction of information flow and the coupling strength between two signals, indicating TE can characterize functional connectivity between two systems.

TE values estimated between the lateral prefrontal cortex and the striatum

We calculated TE values of real LFPs simultaneously recorded in the LPFC and striatum of one monkey performing a stimulus–stimulus association task (Pan et al. 2008). The monkey had to learn two stimulus–stimulus associations in the task, e.g., a visual stimulus A1 was associated with a visual stimulus B1 and a visual stimulus A2 with a visual stimulus B2. Here we briefly introduced the task, more detail information of the task and recording procedures were found in Pan et al. (2008). Each trial of the task started with the onset of a white fixation spot presented

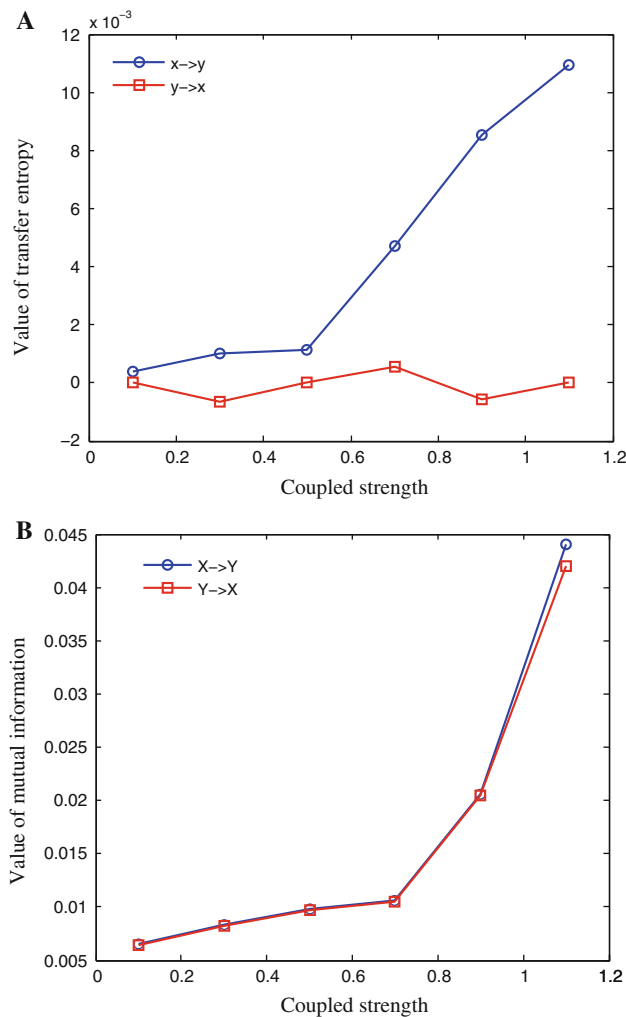


Fig. 5 The TE value and the mutual information as a function of the coupling strength of the two time series. **a** shows the function of TE value against the coupling strength and **b** shows the function of mutual information against the coupling strength. The *blue curves* indicate the TE value or mutual information from the sequence *X* to the sequence *Y*, and *red curves* the TE value or mutual information from the sequence *Y* to the sequence *X*. (Color figure online)

at the center of the monitor. The monkey had to fixate the spot for a random duration (1,000–1,200 ms), and then a sample cue stimulus, for example A1, was presented at the center of the display for 400 ms. After a variable delay period (700–1,200 ms), the fixation spot disappeared and at the same time, the second stimuli, B1 and B2, were presented pseudo-randomly at the left and right positions on the monitor. The subject made a saccade to the target stimulus (e.g. B1) to select it. If the sample stimulus was A2, the monkey chose the stimulus B2 as the correct target.

We inserted one U-probe electrode with 8 channels (Plexon Inc, Texas, USA) into the LPFC, and the other with 8 channels into the striatum to record LFPs simultaneously in the both areas. Figure 6 presents the sample data of LFPs from two U-probes, the 8-channel data recorded

from the LPFC (Fig. 6a) and the 8-channel data from the striatum (Fig. 6b). The data in each channel was aligned with the stimulus onset (indicated by the dashed line), with the duration of 5 s.

We calculate TE values for each pair of channels (one channel from the LPFC and the other from the striatum) in two time epochs: fixation period (900 ms prior to the sample stimulus) and stimulus period (900 ms after the sample stimulus), respectively. It is known that prefrontal neurons directly project to striatal neurons. Outputs of striatal neurons reach the thalamus through the direct and indirectly pathways, and the thalamus feeds back signals to the prefrontal cortex (Alexander et al. 1986). According to such anatomical connections between the prefrontal cortex and the striatum, striatal neurons would receive signals directly from prefrontal neurons, whereas prefrontal neurons would not receive information directly from striatal neurons. Thus we expected that the TE value from the LPFC to the striatum would be larger than that from the striatum to the LPFC.

Figure 7a shows the TE values averaged across pairs of channels in each session and across 10 recording sessions. In the stimulus period, the TE strength from the LPFC to the striatum was significantly greater comparing to that in the opposite direction (two tailed *t* test, $P < 10^{-5}$). The two TE values, however, had no significant difference in the fixation period (two tailed *t* test, $P = 0.3158$). The TE value from the LPFC to the striatum increased in the stimulus period comparing to in the fixation period. The TE value from the striatum to the LPFC had no change between the two periods. Correspondingly, we computed mutual information between the LPFC and striatum based on the same database (Fig. 7b). The value of mutual information significantly increased in the stimulus period than in the fixation period (two-tailed *t* test, $P < 10^{-4}$), indicating mutual information could detect the variation of information transmission in the two periods. But the mutual information from the LPFC to the striatum was completely same as the mutual information from the striatum to the LPFC, indicating the mutual information did not estimate the direction of information transmission. The results suggested that TE not only characterized anatomical projections from one region to another region, but also described dynamical properties of functional connectivity between those areas. The dynamical properties may reflect various interactions between the LPFC and the striatum during different behavior stages.

Discussions

Most approaches used in the investigation of causality in neuroscience are based on interpretation of the Granger Causality definition (Ding et al. 2000; Kaminski et al.

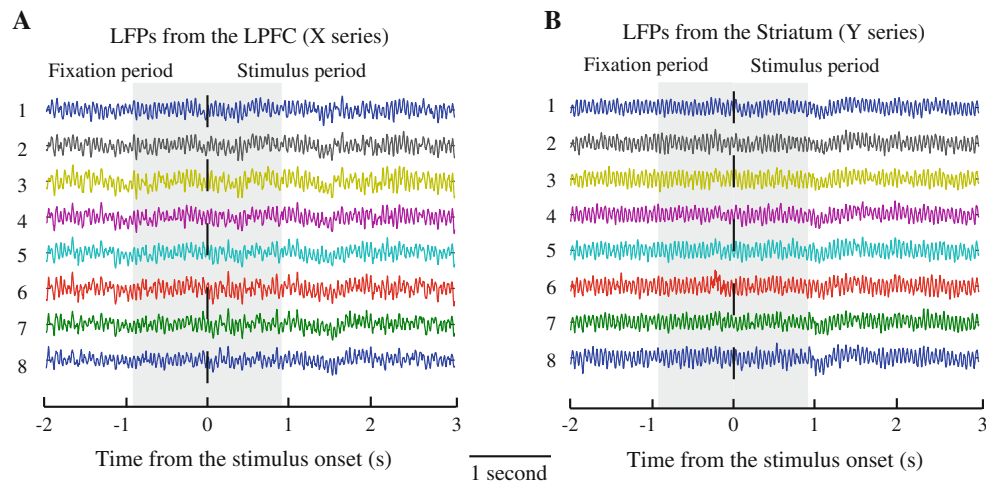


Fig. 6 The raw LFP data of 16 channels recorded in one trial. LFP data recorded in the LPFC by a U-probe with 8 channels is presented in (a) and LFP data recorded in the striatum is shown in (b). LFP data in each channel was segmented and aligned on the stimulus onset (indicated by the *dashed line*). The two *gray areas* (separated by the

dashed line) indicate the two time windows in which LFP data was used to calculate transfer entropy and mutual information. Fixation period is prior to the stimulus onset, and Stimulus period is post the stimulus onset

2001; Seth and Edelman 2007; Seth 2008). One limitation in Granger causality is the *prior* assumption of linear interaction between systems. The mutual information has been widely applied in quantifying the overlap of information content of two neurons (Tsukada et al. 1975, 1976; Borst and Theunissen 1999). However, mutual information contains no directional information (see Figs. 5, 7). The TE, based on information theory, is a very general way to define causality, a way that encompasses both linear and nonlinear relationship. Compared to mutual information, TE characterizes the direction of information flow, which may be an appreciate way to describe the information process in the neural system with layered structure.

In this study, we calculated TE values of two stimulated time series X and Y, and found this method was able to correctly predict the coupling delay time from X to Y (see Figs. 1, 2). We also observed that the TE value increased with the increment of the coupling strength between X and Y (see Fig. 5), and decreased with the noise (see Fig. 4), indicating the TE method was able to estimate the coupling strength of two systems. Moreover, the TE value from X to Y was significantly different from zero only in the case of information was transferred from X to Y. These properties suggested that the TE method quantitatively measured the coupling strength and information flow (or causal interaction) between two systems.

We also calculate the TE values and mutual information between the LPFC and the striatum based on real LFP data. The TE method detected the functional connectivity from the LPFC to the striatum was stronger than the connectivity from the striatum to the LPFC (see Fig. 7), consisting with

the known anatomical evidence between them (Yin and Knowlton 2006). More importantly, we found that the TE value from the LPFC to the striatum was dependent of the monkey's behavior. In the fixation period, the TE value from the LPFC to the striatum did not differ from the TE value in reverse direction. In the stimulus period, the former significantly increased, indicating more LPFC information was sent to the striatum in stimulus period than in the fixation period. Therefore, the TE method had the ability to reveal dynamical changing of functional connectivity between brain regions, which may reflect dynamical interactions in such prefrontal-striatal networks correlated to different behaviors. Mutual information estimated the functional connectivity between the two areas symmetrically, without directional information (see Fig. 7b).

The brain system is a hierarchical organization with forward and feedback connections. A brain region may send its outputs to a higher-order area, and at the same time receives feedback information from the higher-order area, forming a circular circuit. It is usually difficult to distinguish casual relations between those signals temporally. But in the prefrontal-striatal circuits, prefrontal neurons directly project to striatal neurons through monosynaptic connections. On the other hand, striatal neurons have no direct connections to prefrontal neurons. Striatal neurons project in two different pathways: the direct and the indirect pathways (Percheron and Fillion 1991; Kamishina et al. 2008). In the direct pathway, Neurons in the striatum project onto the cells of the Substantia nigra reticuloglobus pallidus interna (SNr-GPi) complex. The SNr-GPi

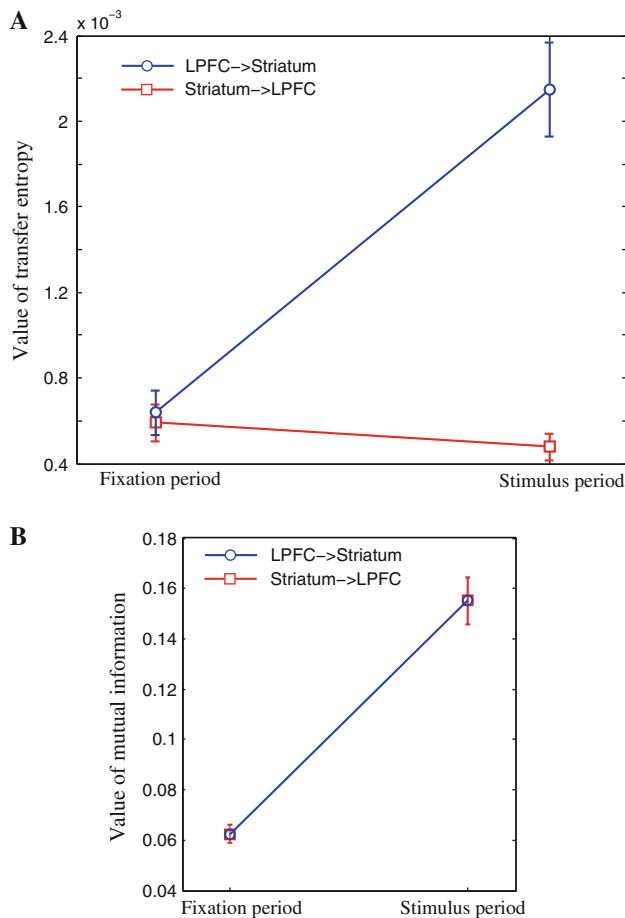


Fig. 7 TE and mutual information between the LPFC and the striatum estimated from real LFP data in the two time epochs. **a** shows TE and **b** shows mutual information. The *blue lines* indicate the TE value or mutual information from the LPFC to the striatum, and the *red lines* represent the TE value or mutual information from the striatum to the LPFC. *Error bars* indicate s.e.m. (Color figure online)

complex projects directly onto the thalamus. Striatal neurons in the indirect pathway project onto the cells of the globus pallidus externa (GPe), which inhibits the subthalamic nucleus (STN). The STN, in turn, projects inputs to the SNr-GPi complex and then to the thalamus. Finally, through the thalamus, the signals are feedback to the prefrontal cortex. In the prefrontal-striatal circuits, the transmission time from the prefrontal cortex to the striatum should be much less than that from the striatum to the prefrontal cortex. We selected the prediction time u as 10 ms, approximating the averaged transmission time from the prefrontal cortex to the striatum (Maurice et al. 1998), to calculate TE of LFP data. Therefore we could find significantly higher TE value from the prefrontal cortex to the striatum than from the reverse direction. It is possible that when we set the prediction time matching the averaged transmission time from the striatum to the prefrontal cortex, TE values from the striatum to the prefrontal cortex

would be larger than TE values from the prefrontal cortex to the striatum. We can temporally separate cause-effect relations in the prefrontal-striatal circuits due to their different transmission times. In two generally coupled systems with similar transmission times, the bidirectional information theory (Marko 1973) could be used to clarify information transmission between them. Further study is necessary to investigate these issues.

Both TE and mutual information were significantly stronger in the stimulus period than in the fixation period (see Fig. 7), suggesting the functional connectivity became more effective from the prefrontal cortex to the striatum in the stimulus period. The changing of the functional connectivity may be due to the increasing of synaptic strength through learning or synchronized firing patterns of two assembled neurons in the prefrontal cortex and striatum. Synaptic weights between prefrontal and striatal neurons would be modified during the learning of the task. After the monkey completed the learning, the weights would keep stable. LFPs were recorded after the monkey got over-trained. We compared TE values in the two time epochs prior to or post the stimulus onset in each trial. It is not likely for prefrontal and striatal neurons to modify their synaptic weights quickly enough across the stimulus onset trial by trial. The increasing TE from the fixation period to the stimulus period may be due to the changing of synchronized firing patterns in the prefrontal cortex and striatum. In the fixation period, neurons in the two areas fire independently, and the temporal correlation between them is low. In the stimulus period, more neurons show synchronized firing and their temporal correlation increases. So the functional connectivity may reflect information transmission or communication between assembled neurons that may synchronize to operate a specific function, and it may depends on the coincidence of arrived EPSPs and IPSPs at post-synaptic neurons (Tsukada et al. 1977; Fingelkurts et al. 2005).

The current TE method estimated causal interactions between two systems in the temporal domain, but not in the frequency domain. Neural signals, like LFP, EEG, contain a very broad frequency spectrum, ranging from a fraction of a Hz to well over 100 Hz, such as θ (4–7 Hz), α (8–13 Hz), β (14–30 Hz) and γ (25–100 Hz) frequency bands (Gu and Liang 2007). Each frequency band and their interactions may play different functional roles in cortical information processing (Siegel et al. 2012); for example, low frequency band is involved in neural communication between different brain regions, while high frequency band is involved in neural communication within a brain region (Buschman and Miller 2007). One important issue that needs call for further investigation is how to apply the TE to analyze causal interactions of two systems in the frequency domain.

Acknowledgments This work was supported by National Foundation of Natural Science of China (No. 11232005) and the Fundamental Research Funds for the Central Universities of China.

References

- Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9:357–381
- Barnett L, Barrett AB, Seth A (2009) Granger causality and transfer entropy are equivalent for Gaussian Variables. *Phys Rev Lett* 103(23): e.238701
- Besserve M, Scholkopf B, Logothetis NK, Panzeri S (2010) Causal relationships between frequency bands of extracellular signals in visual cortex revealed by an information theoretic analysis. *J Comput Neurosci* 29(3):574–576
- Borst A, Theunissen FE (1999) Information theory and neural coding. *Nat Neurosci* 2(11):947–957
- Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–1862
- Cao L (1997) Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D Nonlinear Phenomena* 110(1–2):43–50
- Dan Y, Alonso JM, Usrey MW, Reid RC (1998) Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat Neurosci* 1(6):501–607
- Ding M, Bressler S, Yang W, Liang H (2000) Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Bio Cybern* 83(1):35–45
- Duda RO, Hart PE, Stork DG (2004) *Pattern classification*. Machinery Industry Press, Beijing
- Fingelkurts AA, Fingelkurts AA, Kahkonen S (2005) Functional connectivity in the brain- is it an elusive concept? *Neurosci Biobehav Rev* 28(8):827–836
- Freeman WJ (2007) Definitions of state variables and state space for brain–computer interface Part 2. Extraction and classification of feature vectors. *Cogn Neurodyn* 1(2):85–96
- Garofalo M, Nieuws T, Massobrio P, Martinoia S (2009) Evaluation of the performance of information theory based methods and cross-correlation to estimate the functional connectivity in cortical networks. *PLoS ONE* 4(8):e6482
- Gourévitch B, Eggermont JJ (2007) Evaluating information transfer between auditory cortical neurons. *J Neurophysiol* 97(3):2533–2543
- Granger C (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37(3):424–438
- Gu F, Liang PJ (2007) *Neural information processing* (in Chinese). Beijing University of Technology Press, Beijing
- Hu SQ, Cao Y, Zhang JH, Kong WZ, Yang K, Zhang YB, Li X (2012) More discussions for granger causality and new causality measures. *Cogn Neurodyn* 6(1):33–42
- Ito S, Hansen ME, Heiland R, Lumsdaine A, Litke AM, Beggs JM (2011) Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model. *PLoS ONE* 6(11):e27431
- Kamiński M, Ding M, Truccolo WA, Bressler SL (2001) Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Bio Cybern* 85(2):145–157
- Kamishina H, Yurcisin G, Corwin J, Reep R (2008) Striatal projections from the rat lateral posterior thalamic nucleus. *Brain Res* 1204:24–39
- Kraskov A (2004) Synchronization and interdependence measures and their applications to the Electroencephalogram of epilepsy patients and clustering of data. *NIC Series* 24
- Kraskov A, Stögbauer H, Grassberger P (2004) Estimating mutual information. *Phys Rev E* 69(6): e066138
- Kullback S, Leibler RA (1951) On information and sufficiency. *Ann Mathe Stat* 22(1):79–86
- Lindner M, Vicente R, Priesemann V, Wibral M (2011) Trentool: a matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neurosci*. doi:10.1186/1471-2202-12-119
- Marko H (1973) The bidirectional communication theory—a generalization of information theory. *IEEE Transact commun COM-21* (12): 1345–1351
- Maurice N, Deniau JM, Glowinski J, Thierry AM (1998) Relationships between the prefrontal cortex and the basal ganglia in the rat: physiology of the corticosubthalamic circuits. *J Neurosci* 18(22):9539–9546
- Pan X, Sawa K, Tsuda I, Tsukada M, Samagami M (2008) Reward prediction based on stimulus categorization in primate lateral prefrontal cortex. *Nat Neurosci* 11(6):703–712
- Paninski L (2003) Estimation of entropy and mutual information. *Neural Comput* 15(6):1191–1253
- Percheron G, Filion M (1991) Parallel processing in the basal ganglia: up to a point. *Trends Neurosci* 14(2):55–59
- Schreiber T (2000) Measuring Information Transfer. *Phys Rev Lett* 85(2):461–464
- Seth A (2008) Causal networks in simulated neural systems. *Cogn Neurodyn* 2(1):49–64
- Seth A, Edelman G (2007) Distinguishing causal interactions in neural populations. *Neural Comput* 19(4):910–933
- Shannon CE, Weaver W (1949) *The mathematical theory of information*. University of Illinois Press, Urbana
- Siegel M, Donner TH, Engel AK (2012) Spectral fingerprints of large-scale neuronal interactions. *Nat Rev Neurosci* 13(1):121–134
- Tsukada M, Ishii N, Sato R (1975) Temporal pattern discrimination of impulse sequences in the computer-simulated nerve cells. *Bio Cybern* 17:19–28
- Tsukada M, Ishii N, Sato R (1976) Stochastic automaton models for the temporal pattern discrimination of nerve impulse sequences. *Bio Cybern* 21:121–130
- Tsukada M, Usami H, Sato R (1977) Stochastic automation models for interaction of excitatory and inhibitory impulse sequences in neurons. *Bio Cybern* 27:235–245
- Werner G (2009) Consciousness related neural events viewed as brain state space transitions. *Cogn Neurodyn* 3(1):83–95
- Wiener N (1956) *The theory of prediction*. In: Beckmann EF (ed) *Modern mathematics for the engineer*. McGraw-Hill, New York
- Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7(1):464–476