International Journal of Bifurcation and Chaos, Vol. 22, No. 11 (2012) 1250261 (9 pages)
© World Scientific Publishing Company
DOI: 10.1142/S0218127412502616

SYNCHRONIZATION AS A PROCESS OF SHARING AND TRANSFERRING INFORMATION

ERIK M. BOLLT
Department of Mathematics, Clarkson University,
Box 5815, Potsdam, NY 13699, USA

Received June 16, 2011

Synchronization of chaotic oscillators has become well characterized by errors which shrink relative to a synchronization manifold. This manifold is the identity function in the case of identical systems, or some other slow manifold in the case of generalized synchronizaton in the case of nonidentical components. On the other hand, since many decades beginning with the Smale horseshoe, chaotic oscillators can be well understood in terms of symbolic dynamics as information producing processes. We study here the synchronization of a pair of chaotic oscillators as a process for sharing information bearing bits transferred between each other, by measuring the transfer entropy tracked as the global system transitions to the synchronization state. Further, we present for the first time the notion of transfer entropy in the measure theoretic setting of transfer operators.

Keywords: Dynamical systems; chaos; networks; synchronization; information theory; entropy; transfer entropy.

1. Introduction

The phenomena of synchronization has been found in various aspects of nature and science [Strogatz, 1994]. Its applications have ranged widely from biology [Strogatz & Stewart, 1993; Golubitsky *et al.*, 1999] to mathematical epidemiology [He & Stone, 2003], and chaotic oscillators [Pecora & Carroll, 1990], to communicational devices in engineering [Cuomo & Oppenheim, 1993], etc. Generally, the analysis of chaotic synchronization has followed a discussion of the synchronization manifold, which may be the identity function for identical oscillators [Pecora & Carroll, 1998], or some perturbation thereof for nonidentical oscillators [Sun *et al.*, 2009], often by some form of master stability function analysis.

In the perspective of information theory and symbolic dynamics, it can be understood that chaotic oscillators are essentially information bearing sources producing symbols at each iteration in time, with the entropy being descriptive of the rate of information production [Robinson & Robinson, 1999; Bollt, 2003]. In this vein, we will study here a different perspective on synchronization than the normal analytic one. We will consider coupled oscillators as sharing information, and the process of synchronization as being one where the shared information is an entrainment of the entropy production. To understand this sharing of information, we will resort to the transfer entropy of Schreiber [2000]. In this perspective, when oscillators synchronize, it can be understood that they must be sharing symbols in order that they may each express the same symbolic dynamics. Furthermore, depending on the degree of co-coupling, or master-slave coupling or somewhere in between, the directionality of the information flow can be described by the transfer entropy. For the sake of a related work, we wish to point the reader to a study of anticipating synchronization which has a transfer entropy perspective while studying the appropriate scale necessary to infer directionality [Hahs & Pethel, 2011].

In our presentation, we choose the following skew tent map system to use as a coupling element [Hasler & Maistrenko, 1997] which is of a full folding form [Billings & Bollt, 2001]

$$f_a(x) = \begin{Bmatrix} \dfrac{x}{a} & \text{if } 0 \le x \le a \\ \dfrac{1-x}{1-a} & \text{if } a \le x \le 1 \end{Bmatrix}, \tag{1}$$

that we couple in the following nonlinear form [Hasler & Maistrenko, 1997],

$$\begin{aligned} \begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} &= \mathbf{G}\begin{pmatrix} x_n \\ y_n \end{pmatrix} \\ &= \begin{pmatrix} f_{a_1}(x_n) + \delta(y_n - x_n) \\ f_{a_2}(y_n) + \epsilon(x_n - y_n) \end{pmatrix}. \end{aligned} \tag{2}$$

We see that written in this form, if $a_1 = a_2$ and $\epsilon = 0$ but $\delta > 0$ we have a master-slave system of identical systems as illustrated in Fig. 1 where we see a stable synchronized identity manifold where error decreases exponentially to zero. On the other hand, if $\epsilon = \delta$ but $a_1 \neq a_2$ we can study symmetrically coupled but nonidentical systems in Figs. 2 and 3, where the identity manifold is not exponentially stable but is apparently a Lyapunov stable manifold as the error, $\text{error}(n) = |x(n) - y(n)|$ remains small for both scenarios shown in the figures, $a_1 = 0.63$ but $a_2 = 0.65$ and $a_2 = 0.7$ respectively, with progressively larger but stable errors. Our presentation here will be designed to introduce the perspective of transfer entropy to understand the process of synchronization in terms of information flow, and from this perspective to gain not only an idea of when

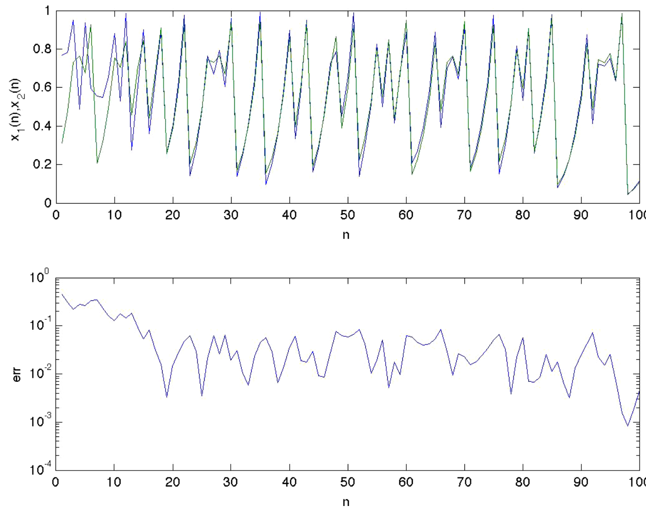

Fig. 2. A nonlinearly coupled skew tent map system, Eq. (2), of nonidentical oscillators, $a_1 = 0.63, a_2 = 0.65$ and master-slave configuration, $\delta = 0.6$, $\epsilon = 0.0$. Note (above) how the signals *approximately* entrain and (below) the error, $\text{error}(n) = |x(n) - y(n)|$ decreases close to zero, where it remains *close* to an identity manifold, $x = y$ where it is stable in a Lyapunov stability sense.

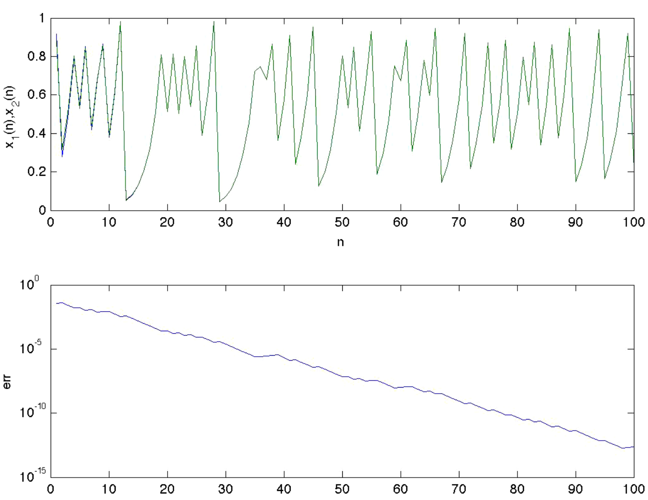

Fig. 1. In a nonlinearly coupled skew tent map system, Eq. (2), of identical oscillators, $a_1 = a_2 = 0.63$ and master-slave configuration, $\delta = 0.6$, $\epsilon = 0.0$ (parameters as in [Hasler & Maistrenko, 1997]). Note (above) how the signals entrain and (below) the error, $\text{error}(n) = |x(n) - y(n)|$ decreases exponentially.

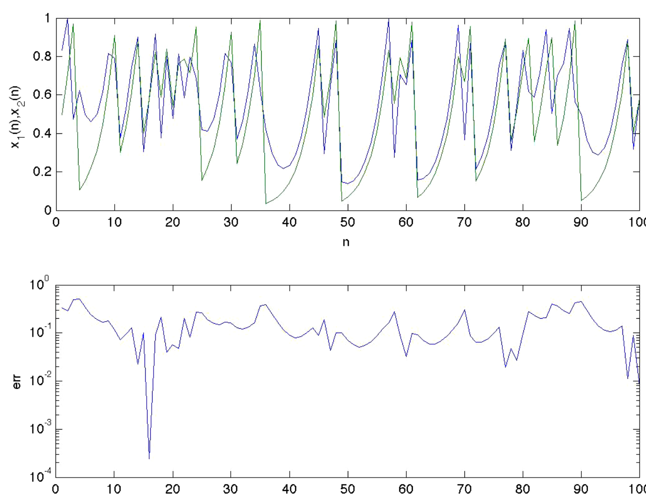

Fig. 3. A nonlinearly coupled skew tent map system, Eq. (2), of nonidentical oscillators, $a_1 = 0.63, a_2 = 0.7$ and master-slave configuration, $\delta = 0.6$, $\epsilon = 0.0$. As in Fig. 2, note (above) how the signals *approximately* entrain and (below) the error, $\text{error}(n) = |x(n) - y(n)|$ decreases close to zero, but not as small as in Fig. 2.

oscillators synchronize but perhaps if one or the other is acting as a master or a slave. Furthermore, the perspective is distinct from a master stability formalism.

2. Information Flow and Transfer Entropy

A natural question in measurable dynamical systems is to ask which parts of a partitioned dynamical system influence other parts of the system. Detecting dependencies between variables is a general statistical question and in a dynamical systems context this relates to questions of causality. There are many ways one may interpret and subsequently computationally address dependency. For example, familiar linear methods such as correlation have some relevance, and these methods are very popular especially for the simplicity of application [Kantz *et al.*, 1997]. A popular method is to compute mutual information, $I(X_1; X_2)$ [Cover *et al.*, 1991],[1]

$$dI(X_1; X_2) = \sum_{x_1, x_2} p(x_1, x_2) \log \frac{p(x_1, x_2)}{p_1(x_1) p_2(x_2)} \tag{3}$$

as a method to consider dynamical influence such as used in [Donges *et al.*, 2009] in the context of global weather events. However, both correlation and mutual information more so address overlap of states rather than information flow and therefore time dependencies.

The transfer entropy $T_{J \to I}$ was recently developed by Schreiber [2000] to be a statistical measure of information flow, with respect to time, between states of a partitioned phase space in a dynamical system to other states in a dynamical system. Unlike other methods that simply consider common histories, transfer entropy explicitly computes information exchange in a dynamical signal. Here we will review the ideas behind transfer entropy as a measurement of causality in a time evolving system, and then we will show how this quantity can be computed using estimates of the Frobenius–Perron transfer operator by carefully masking the resulting matrices. We follow here the notation of our book in progress [Bollt *et al.*, 2013]. Note that in the case of Gaussian noise, it has been shown that Granger causality measure has been shown to be equivalent to transfer entropy [Barnett *et al.*, 2009].

2.1. *Definitions and interpretations of transfer entropy*

To discuss transfer entropy, suppose that we have a partitioned dynamical system on a skew product space $X \times Y$,

$$F : X \times Y \to X \times Y. \tag{4}$$

This notation of a single dynamical system with phase space written as a skew product space allows a broad application as we highlight in the examples and help to clarify the transfer of entropy between the X and Y states. For now, we will further write this system as if it is a two-coupled dynamical system having x and y parts describing the action on each component and perhaps with coupling between components,

$$F(x, y) = (F_x(x, y), F_y(x, y)), \tag{5}$$

where,

$$\begin{aligned} T_x &: X \times Y \to X \\ &\qquad x_n \mapsto x_{n+1} = T_x(x_n, y_n) \end{aligned} \tag{6}$$

and likewise,

$$\begin{aligned} T_y &: X \times Y \to Y \\ &\qquad y_n \mapsto y_{n+1} = T_y(x_n, y_n). \end{aligned} \tag{7}$$

This notation allows that $x \in X$ and $y \in Y$ may each be vector (multivariate) quantities and even of dimensions different from each other. See Fig. 4.

Let,

$$x_n^{(k)} = (x_n, x_{n-1}, x_{n-2}, \ldots, x_{n-k+1}), \tag{8}$$

be the measurements of a dynamical system T_x, at times,

$$t^{(k)} = (t_n, t_{n-1}, t_{n-2}, \ldots, t_{n-k+1}), \tag{9}$$

sequentially. In this notation, the space X is partitioned into states $\{x\}$ and hence x_n denotes the measured state at time t_n. Note that we have chosen here not to index in any way the partition $\{x\}$, which may be some numerical grid as shown in Fig. 4, since subindices are already being used to denote time, and super indices denote time-depth of the sequence discussed, so an index to denote space would be a bit of notation overload. We may denote

[1] It is useful to point out at this stage that $p_1(x_1)$ and $p_2(x_2)$ are the marginal distributions of $p(x_1, x_2)$; $p_1(x_1) = \sum_{x_2} p(x_1, x_2)$ and likewise, $p_2(x_2) = \sum_{x_1} p(x_1, x_2)$.

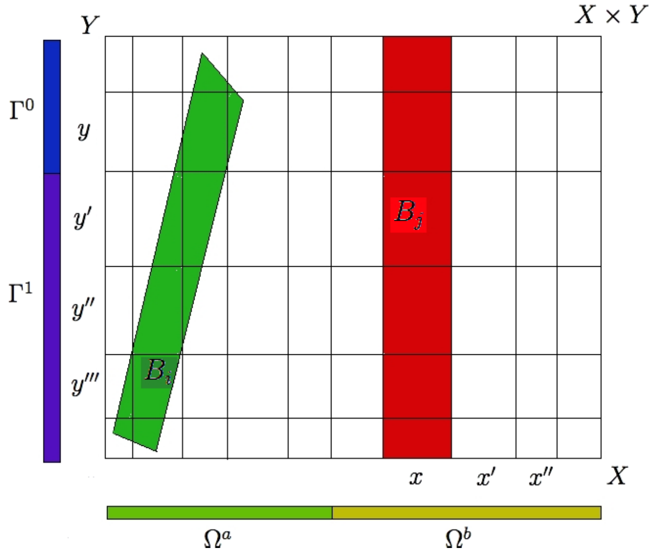

Fig. 4. In a skew product space $X \times Y$, to discuss transfer entropy between states $\{x\}$ a partition of X and states $\{y\}$ of Y, some of which are illustrated as x, x', x'' and y, y', y'', y''', a coarser partition $\{\Omega^a, \Omega^b\}$ of X in symbols a and b and likewise $\{\Gamma^0, \Gamma^1\}$ of Y in symbols 0 and 1 are also illustrated.

simply x, x' and x'' to distinguish states where needed. Likewise, $y_n^{(k)}$ denotes sequential measurements of y at times $t^{(k)}$, and also Y may be partitioned into states $\{y\}$ as characterized in Fig. 4.

The main idea leading to transfer entropy [Schreiber, 2000] is to measure the deviation from the Markov property, which would presume,

$$p(x_{n+1} \,|\, x_n^{(k)}) = p(x_{n+1} \,|\, x_n^{(k)}, y_n^{(l)}), \tag{10}$$

that the state $(x_{n+1} \,|\, x_n^{(k)})$ does not include dependency on $y_n^{(l)}$. When there is a departure from this Markovian assumption, the suggestion is that there is no information flow as conditional dependency in time from y to x. The measurement of this deviation between these two distributions is by a conditional Kullback–Leibler divergence which we will build toward in the following.

The joint entropy [Cover *et al.*, 1991] of a sequence of measurements is,

$$H(x_n^{(k)}) = -\sum_{x_n^{(k)}} p(x_n^{(k)}) \log p(x_n^{(k)}). \tag{11}$$

A conditional entropy [Cover *et al.*, 1991],

$$\begin{aligned} H(x_{n+1} \,|\, x_n^{(k)}) &= -\sum p(x_{n+1}, x_n^{(k)}) \log p(x_{n+1} \,|\, x_n^{(k)}) \\ &= H(x_{n+1}, x_n^{(k)}) - H(x_n^{(k)}) \\ &= H(x_{n+1}^{(k+1)}) - H(x_n^{(k)}), \end{aligned} \tag{12}$$

is approximately an *entropy rate* which as it is written quantifies the amount of new information that a new measurement of x_{n+1} allows following the k-prior measurements, $x_n^{(k)}$. Note that the second equality follows the probability chain rule,

$$p(x_{n+1} \,|\, x_n^{(k)}) = \frac{p(x_{n+1}^{(k+1)})}{p(x_n^{(k)})} \tag{13}$$

and the last equality from the notational convention for writing the states,

$$\begin{aligned} (x_{n+1}, x_n^{(k)}) &= (x_{n+1}, x_n, x_{n-1}, \ldots, x_{n-k+1}) \\ &= (x_{n+1}^{(k+1)}). \end{aligned} \tag{14}$$

Transfer entropy is defined in terms of a Kullback–Leibler divergence, $D_{KL}(p_1 \| p_2)$ [Cover *et al.*, 1991] but of conditional probabilities,[2]

$$D_{KL}(p_1(A \,|\, B) \| p_2(A \,|\, B)) = \sum_{a,b} p_1(a, b) \log \frac{p_1(a \,|\, b)}{p_2(a \,|\, b)}, \tag{15}$$

but for states specifically designed to highlight transfer of entropy between the states X to Y (or vice versa Y to X) of a dynamical system written as skew product, Eq. (4). Define [Schreiber, 2000],

$$T_{x \to y} = \sum p(x_{n+1}, x_n^{(k)}, y_n^{(l)}) \log \frac{p(x_{n+1} \,|\, x_n^{(k)}, y_n^{(l)})}{p(x_{n+1} \,|\, x_n^{(k)})}, \tag{16}$$

which we see may be equivalently written as a difference of entropy rates — like conditional entropies,

$$T_{y \to x} = H(x_{n+1} \,|\, x_n^{(l)}) - H(x_{n+1} \,|\, x_n^{(l)}, y_n^{(k)}). \tag{17}$$

[2]Recall that the Kullback–Leibler divergence of a single random variable A with probability distribution is an error-like quantity describing the entropy difference between the true entropy using the correct coding model $\log p_1(A)$ versus a coding model $\log p_2(A)$ with a model distribution $p_2(A)$ of A. This difference of conditional Kullback–Leibler divergence is a direct application for conditional probability $p_1(A \,|\, B)$ with a model $p_2(A \,|\, B)$.

This may be a most useful form for computation, but for interpretation, a useful form is in terms of a conditional Kullback–Leibler divergence,

$$T_{y\to x} = D_{KL}(p(x_{n+1}\,|\,x_n^{(k)}, y_n^{(l)})\|p(x_{n+1}\,|\,x_n^{(k)})), \tag{18}$$

found by putting together Eqs. (15) and (16). In this form, as already noted in Eq. (10), the transfer entropy has the interpretation as a measurement of the deviation from the Markov property, which would be the truth of Eq. (10). The state $(x_{n+1}\,|\,x_n^{(k)})$ does not include the dependency on $y_n^{(l)}$ suggesting that there is no information flow as a conditional dependency in time from y to x causing an influence on transition probabilities of x. In this sense, the conditional Kullback–Leibler divergence Eq. (18) describes the deviation of the information content from the Markovian assumption. In this sense, $T_{y\to x}$ describes an information flow from the subsystem y to subsystem x. Likewise, and asymmetrically,

$$T_{x\to y} = H(y_{n+1}\,|\,y_n^{(l)}) - H(y_{n+1}\,|\,x_n^{(l)}, y_n^{(k)}), \tag{19}$$

and it is immediate to note that generally,

$$T_{x\to y} \neq T_{y\to x}. \tag{20}$$

This is not a surprise both on the grounds that it has already been stated that Kullback–Leibler divergence is not symmetric, but also there is no prior expectation that influences should be equal.

2.2. *Computation of transfer entropy by constrained Frobenius–Perron kernel*

The main issue in discussing transfer entropy is to ask about information moving from one part of a partitioned phase space to another part of the phase space as time advances. The key to computation is joint probabilities and conditional probabilities as they appear in Eqs. (17) and (19). There are two major ways we may make estimates of these probabilities, but both involve course-graining the states. A direct application of formulas [Eqs. (11) and (12)] and likewise for the joint conditional entropy to Eq. (16) allows,

$$\begin{aligned} T_{y\to x} &= [H(x_{n+1}, x_n) - H(x_n)] \\ &\quad - [H(x_{n+1}, x_n, y_n) - H(x_n, y_n)], \end{aligned} \tag{21}$$

which serves as a useful method of direct computation. For the sake of interpretation, here we will discuss the computation of transfer entropy by computing associated conditional transition probabilities from a matrix estimated from an appropriate operator. Starting with a constrained Frobenius–Perron operator [Szabo & Tel, 1994; Bollt *et al.*, 2002], we can develop an approximation of the necessary entropies by simple matrix operations.

For the sake of notation consistent with the skew product discussion of transfer entropy in Eq. (4), we will denote,

$$z = (x, y) \in X \times Y := Z. \tag{22}$$

Using this notation, recall that the general form of a Frobenius–Perron operator may be written as an integration against a kernel $k(z, \bar{z})$,

$$P_f[p(z)] = \int_Z k(z, \bar{z})\rho(\bar{z})d\bar{z}, \tag{23}$$

covers the deterministic case as a generalized function,

$$k(z, \bar{z}) = \delta(z - F(\bar{z})), \tag{24}$$

or the stochastic cases of additive/multiplicative [Santitissadeekorn & Bollt, 2007] when,

$$k(z, \bar{z}) = g((z - F(y))S^{-1}(\bar{z}))|J|. \tag{25}$$

Considering transfer entropy $T_{x\to y}$ as computed by the formula Eq. (16), we must produce estimates of the probabilities and conditional probabilities,

$$\begin{gathered} p(x_{n+1}, x_n^{(k)}, y_n^{(l)}), p(x_{n+1}\,|\,x_n^{(k)}, y_n^{(l)}) \quad \text{and} \\ p(x_{n+1}\,|\,x_n^{(k)}), \end{gathered} \tag{26}$$

over all states x in the partitioned X and y in Y. We will focus on $p(x_{n+1}, x_n^{(k)}, y_n^{(l)})$, for the sake of discussion. Since $z = (x, y)$, we write,

$$\begin{aligned} p(x_{n+1}, x_n, y_n) &= \sum_{y_{n+1}\in\{y\}} p(x_{n+1}, y_{n+1}, x_n, y_n) \\ &\equiv \sum_{y_{n+1}\in\{y\}} p(z_{n+1}, z_n) \end{aligned} \tag{27}$$

and we have chosen the simplest but most common delay scenario $k = l = 1$. We remind that $\{y\}$ denotes the partition of Y and $y_n \in \{y\}$ here denotes the state in that partition at time t_n, rather than a singleton in Y. In this case, each $p(z_{n+1}, z_n)$ in Eq. (27) can be interpreted as a matrix entry of an Ulam–Galerkin matrix [Ding & Zhou, 1996; Bollt, 2000] by projection using characteristic functions supported over the chosen partition,

$$\{x\} \times \{y\} := \{z\}. \tag{28}$$

Recall that an Ulam–Galerkin matrix approximation of the Frobenius–Perron operator is a finite rank projection onto basis functions which are generally basis functions chosen to be characteristic functions supported over boxes. Specifically, rewrite as a conditional probability,

$$\begin{aligned} A_{i,j} &= \frac{m(B_i \cap F^{-1}(B_j))}{m(B_i)} \\ &= p(z_n \in B_i \,|\, z_{n-1} \in B_j). \end{aligned} \tag{29}$$

In this equation, we have written the partition $\{z\}$ in the notation $\{z\} = \{B_i\}$ as suggested in Fig. 4, whereas in this section there is some abuse of notation as z_n may denote a partition element of $\{z\}$ at time $t = t_n$ or a singleton point as the context may merit.

Choosing characteristic function $\{\chi_{B_i}(z)\}$,

$$\begin{aligned} A_{i,j} &= \langle P_F[\chi_{B_j}(z)], \chi_{B_i}(z)\rangle \\ &= \int_{B_i}\int_{B_j} k(z,\bar{z})\rho(\bar{z})d\bar{z}dz, \end{aligned} \tag{30}$$

these transition probabilities are in terms of a constrained Frobenius–Perron operator which we write, $P_{F,B_j}[\chi_{B_j}(z)]$. Due to integration over a restricted domain defined as,

$$P_{F,B_j}[\rho] = \int_{B_j} k(z,\bar{z})\rho(\bar{z})d\bar{z}. \tag{31}$$

Consider then an indicator vector,

$$\begin{aligned} [v_{B_j}] = 1 &\quad \text{in the element corresponding to } B_j \\ &\qquad \text{and} \\ 0 &\quad \text{in all other positions.} \end{aligned} \tag{32}$$

which can be understood by inspection of Fig. 4. As such, we can write,

$$\begin{aligned} [v_x]_k = 1 &\quad \text{if } k \text{ corresponds to } \Pi_x(B_k) = x, \\ &\qquad \text{and} \\ 0 &\quad \text{otherwise,} \end{aligned} \tag{33}$$

where $\Pi_x : Z \to X$ is a projection function. The vector $[v_x]$ indicates 1 exactly in all those rectangles B_j which, for example, as illustrated in Fig. 4, correspond to the red strip projecting to an element $x \in \{x\}$. Likewise, let,

$$\begin{aligned} [v_y]_k = 1 &\quad \text{if } k \text{ corresponds to } \Pi_y(B_k) = y, \\ &\qquad \text{and} \\ 0 &\quad \text{otherwise,} \end{aligned} \tag{34}$$

with $\Pi_y : Z \to Y$. Counting, if $\{x\}$ has N_x elements, and $\{y\}$ has N_y elements, then $N = N_x N_y$ rectangles are indicated in the partition $\{z\} = \{x\} \times \{y\}$. Then for given element x, indicator vectors of the form $[v_x]$ are $N \times 1$ labeling each $\{B_i\}_{i=1}^N$ in a strip.

We now rewrite Eq. (27) as the following Ulam–Galerkin matrix approximation of the constrained Frobenius–Perron operator, Eq. (30),

$$p(x_{n+1}, z_n) = A \cdot \mathrm{diag}([v_{x_n}]) \cdot p, \tag{35}$$

which follows from the conditional probability chain rule.[3] We must interpret the probability vector $p = p(z_n)$, and therefore the statement $p(x_{n+1}, z_n)$ is a vector output across values z_n. On the one hand, it may be understood to be the usual stationary distribution vector that can be found by a histogram of a sample orbit or by A as the dominant eigenvector; it is feasible however that $p_i = 0$ for some i thus upsetting any division to compute conditionals. On the other hand, in the spirit of developing the Ulam–Galerkin matrix estimate of A in terms of the one step action of the map on a uniform measure, Eq. (29), we will interpret $p_i = m(B_i)$. Such estimates are in the spirit of coarse-graining previously used in the topic of developing probabilities to estimate entropies [Hlavácková-Schindler *et al.*, 2007; Kantz & Schreiber, 2004; Cover *et al.*, 1991]. We interpret $\mathrm{diag}([v_{x_n}])$ as the square matrix whose diagonal entries are from the vector $[v_{x_n}]$ that serves as a masking matrix to constrain the Frobenius–Perron operator just as Eq. (31), so that $A \cdot \mathrm{diag}([v_{x_n}])$ approximate P_{F,x_n}.

The other two probabilities in Eq. (26) are somewhat simpler to derive from A and p

$$p(x_{n+1} \,|\, x_n) = \frac{p(x_{n+1}, x_n)}{p(x_n)}, \tag{36}$$

[3] $P(a \,|\, b) = P(a,b)/P(b)$.

which exists for those x_n such that $p(x_n) \neq 0$, and where,

$$p(x_{n+1}, x_n) = \sum_{y_n \in \{y\}} p(x_{n+1}, z_n), \tag{37}$$

since $p(x_{n+1}, z_n) = p(x_{n+1}, x_n, y_n)$ is summed for each y_n, coming from Eq. (35). Also,

$$p(x_n) = \mathrm{diag}([v_{x_n}]) \cdot p. \tag{38}$$

Finally in Eq. (26),

$$p(x_{n+1} \,|\, z_n) = \frac{p(x_{n+1}, z_n)}{p(z_n)}, \tag{39}$$

which exists for those rectangles z_n such that $p(z_n) \neq 0$. Again $p(x_{n+1}, z_n)$ comes from Eq. (35). Therefore,

$$p(x_{n+1} \,|\, z_n) = \frac{p(x_{n+1}, z_n)}{p(z_n)}, \tag{40}$$

or,

$$p(x_{n+1}, z_n) = A \cdot \mathrm{diag}([v_{x_n}]) \cdot \frac{p}{p}. \tag{41}$$

Substitution of Eqs. (35), (36), (39) into Eq. (16) provides an estimate for $T_{x\to y}$. However, the states in the fine grid may not be those desired to define symbolization of information states. If the transformation is Markov on the skew product space, then representation of A in Eq. (30) is exact. The partition $\{z\}$ serves as a symbolization which in projections by Π_x and Π_y are the grids $\{x\}$ and $\{y\}$ respectively. It may be more useful to consider information transfer in terms of a coarser statement of states. For example, see Fig. 4 where we represent a partition Ω and Γ of X and Y, respectively. For convenience of presentation we represent two states in each partition,

$$\Omega = \{\Omega^a, \Omega^b\} \quad \text{and} \quad \Gamma = \{\Gamma^0, \Gamma^1\}. \tag{42}$$

In this case, then the estimates of all of the several probabilities can be summed in a manner just discussed above. Then the transfer entropy $T_{x\to y}$ becomes in terms of the states of the coarse partitions. The question of how a coarse partition may represent the transfer entropy of a system relative to what would be computed with a finer partition has been discussed in [Hahs & Pethel, 2011] with the surprising result that the direction of information flow can be effectively measured as not just a poor estimate by the coarse partition, but possibly even of the wrong sign.

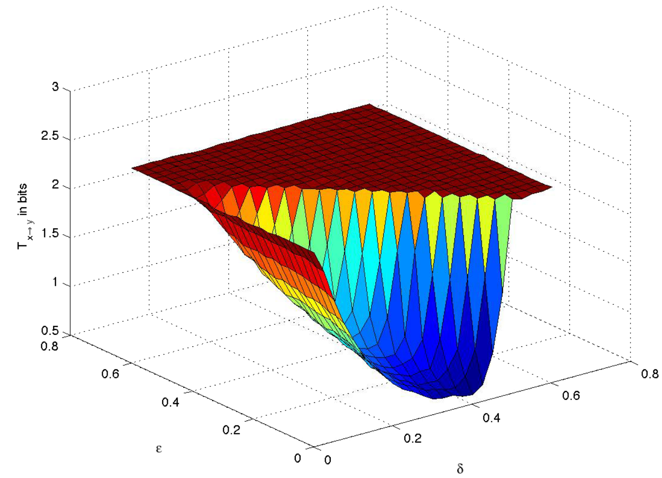

Fig. 5. Transfer entropy, $T_{x\to y}$ measured in bits, of the system Eq. (2), in the identical parameter scenario $a_1 = a_2 = 0.63$ which often results in synchronization depending on the coupling parameters swept, $0 \leq \delta \leq 0.8$ and $0 \leq \epsilon \leq 0.8$ as shown. Contrast to $T_{y\to x}$ as shown in Fig. 6 where the transfer entropy clearly has an opposite phase relative to the coupling parameters (ϵ, δ).

2.3. *Results of information flow due to synchrony*

Now considering the system of coupled skew tent maps, Eq. (2), with coupling resulting in various identical and nonidentical synchronization scenarios as illustrated in Figs. 1–3, we will now analyze the information transfer across a study of both parameter matches and mismatches and across various coupling strengths and directionalities. In Figs. 5 and 6,

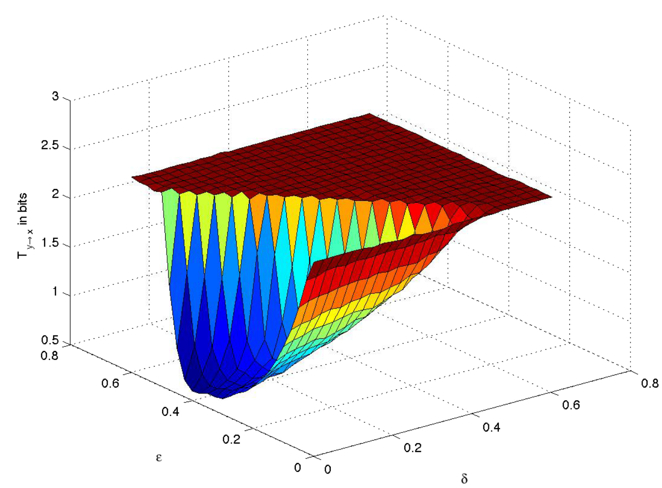

Fig. 6. Transfer entropy, $T_{y\to x}$ measured in bits, of the system Eq. (2), in the identical parameter scenario $a_1 = a_2 = 0.63$ which often results in synchronization depending on the coupling parameters swept, $0 \leq \delta \leq 0.8$ and $0 \leq \epsilon \leq 0.8$ as shown. Compare to $T_{x\to y}$ as shown in Fig. 5.

we see the results of transfer entropy, $T_{x\to y}$ and $T_{y\to x}$ respectively in the scenario of identical oscillators $a_1 = a_2 = 0.63$ for coupling parameters being swept $0 \le \delta \le 0.8$ and $0 \le \epsilon \le 0.8$. We see that due to the symmetry of the form of the coupled systems, Eq. (2), the mode of synchronization is opposite to what is expected. When $T_{x\to y}$ is relatively larger than $T_{y\to x}$, the interpretation is that relatively more information flows from the x system to the y system, and vice versa. This source of communication is due to the coupling of the formulation of synchronization. Large changes in this quantity signals the sharing of information leading to synchronization.

In the asymmetric case, $0.55 \le a_1, a_2 \le 0.65$ we show a master-slave coupling $\epsilon = 0$, $\delta = 0.6$ in Fig. 7 and compare with Figs. 1–3. In the master-slave scenario chosen, the x oscillator is driving the xy oscillator. As such, the x oscillator sends its states in the form of bits to the y oscillator as should be measured as $T_{x\to y} > T_{y\to x}$ when synchronizing and more so when a great deal of information "effort" is required to maintain synchronization. This is interpreted as seen in Fig. 7 in that when the oscillators are identical, $a_1 = a_2$ shown on the diagonal, the transfer entropy difference $T_{x\to y} > T_{y\to x}$ is smallest since the synchronization requires the smallest exchange of information once started. In contrast, $T_{x\to y} > T_{y\to x}$ is the largest when the oscillators are most dissimilar, and we see in Fig. 5 how "strained"

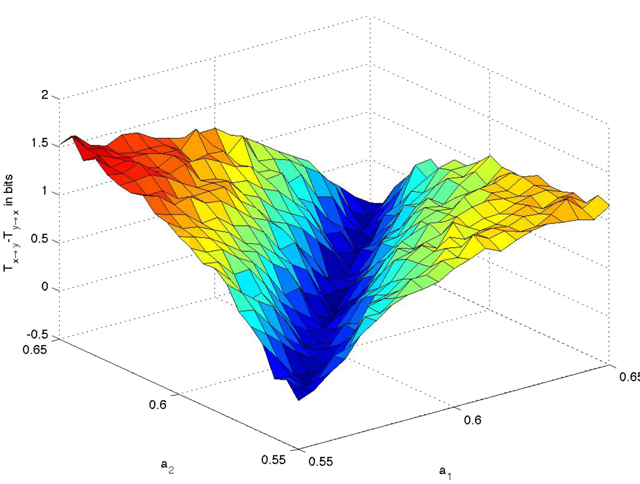

Fig. 7. Transfer entropy difference, $T_{y\to x} - T_{y\to x}$ measured in bits, of the system Eq. (2), in the nonidentical parameter scenario sweep, $0.55 \le a_1, a_2 \le 0.65$, and master-slave coupling $\epsilon = 0$, $\delta = 0.6$. Compare to $T_{x\to y}$ as shown in Fig. 5. Contrast to $T_{y\to x}$ as shown in Fig. 6 where the transfer entropy clearly has an opposite phase relative to the coupling parameters (ϵ, δ). Also compare to Figs. 1–3.

the synchronization can be since the error cannot go to zero as the oscillators are only loosely bound.

3. Conclusion

We have discussed here the concept of synchronization of both identical and nonidentical systems with various directionalities and strengths of couplings in a new language of information using Transfer entropy. We have developed the tool of transfer entropy a bit further than previously presented by the interpretation within the formalism of transfer operators by the Frobenius–Perron operator. We have found that transfer entropy can make a useful tool for interpreting the synchrony that develops between coupled oscillators. An obvious next direction for this work is to apply these methods to explore synchrony within complex networks of oscillators, but now within the language of information theory to discuss information flowing within a complex network.

Acknowledgment

This work has been supported by the National Science Foundation DMS-0708083 and by the Office of Naval Research GRANT00539561.

References

Barnett, L., Barrett, A. & Seth, A. [2009] "Granger causality and transfer entropy are equivalent for Gaussian variables," *Phys. Rev. Lett.* **103**, 238701.

Billings, L. & Bollt, E. [2001] "Probability density functions of some skew tent maps," *Chaos Solit. Fract.* **12**, 365–376.

Bollt, E. [2000] "Controlling chaos and the inverse Frobenius–Perron problem: Global stabilization of arbitrary invariant measures," *Int. J. Bifurcation and Chaos* **10**, 1033–1050.

Bollt, E., Billings, L. & Schwartz, I. [2002] "A manifold independent approach to understanding transport in stochastic dynamical systems," *Physica D* **173**, 153–177.

Bollt, E. [2003] "Review of chaos communication by feedback control of symbolic dynamics," *Int. J. Bifurcation and Chaos* **13**, 269–285.

Bollt, E. & Santitissadeekorn, N. [2011] *Applied and Computational Measurable Dynamics* (SIAM).

Cover, T., Thomas, J., Wiley, J. *et al.* [1991] *Elements of Information Theory*, Vol. 306 (Wiley Online Library).

Cuomo, K. & Oppenheim, A. [1993] "Circuit implementation of synchronized chaos with applications to communications," *Phys. Rev. Lett.* **71**, 65–68.

Ding, J. & Zhou, A. [1996] "Finite approximations of Frobenius–Perron operators. A solution of Ulam's conjecture to multi-dimensional transformations," *Physica D* **92**, 61–68.

Donges, J., Zou, Y., Marwan, N. & Kurths, J. [2009] "The backbone of the climate network," *Europhys. Lett.* **87**, 48007.

Golubitsky, M., Stewart, I., Buono, P. & Collins, J. [1999] "Symmetry in locomotor central pattern generators and animal gaits," *Nature* **401**, 693–695.

Hahs, D. & Pethel, S. [2011] "Distinguishing anticipation from causality: Anticipatory bias in the estimation of information flow," *Phys. Rev. Lett.* **107**, 128701.

Hasler, M. & Maistrenko, Y. [1997] "An introduction to the synchronization of chaotic systems: Coupled skew tent maps," *IEEE Trans. Circuits Syst.-I*: *Fund. Th. Appl.* **44**, 856–866.

He, D. & Stone, L. [2003] "Spatio-temporal synchronization of recurrent epidemics," *Proc. Roy. Soc. London. Series B*: *Biol. Sci.* **270**, 1519.

Hlavácková-Schindler, K., Palus, M., Vejmelka, M. & Bhattacharya, J. [2007] "Causality detection based on information-theoretic approaches in time series analysis," *Phys. Rep.* **441**, 1–46.

Kantz, H., Schreiber, T. & Mackay, R. [1997] *Nonlinear Time Series Analysis*, Vol. 2000 (Cambridge University Press, Cambridge).

Kantz, H. & Schreiber, T. [2004] *Nonlinear Time Series Analysis* (Cambridge University Press).

Pecora, L. & Carroll, T. [1990] "Synchronization in chaotic systems," *Phys. Rev. Lett.* **64**, 821–824.

Pecora, L. & Carroll, T. [1998] "Master stability functions for synchronized coupled systems," *Phys. Rev. Lett.* **80**, 2109–2112.

Robinson, R. & Robinson, C. [1999] *Dynamical Systems*: *Stability*, *Symbolic Dynamics*, *and Chaos* (CRC).

Santitissadeekorn, N. & Bollt, E. [2007] "The infinitesimal operator for the semigroup of the Frobenius–Perron operator from image sequence data: Vector fields and transport barriers from movies," *Chaos* **17**, 023126.

Schreiber, T. [2000] "Measuring information transfer," *Phys. Rev. Lett.* **85**, 461–464.

Strogatz, S. & Stewart, I. [1993] "Coupled oscillators and biological synchronization," *Scient. Amer.* **269**, 102–109.

Strogatz, S. [1994] *Nonlinear Dynamics and Chaos*: *With Applications to Physics*, *Biology*, *Chemistry*, *and Engineering* (Westview Press).

Sun, J., Bollt, E. & Nishikawa, T. [2009] "Master stability functions for coupled nearly identical dynamical systems," *Europhys. Lett.* **85**, 60011.

Szabo, K. & Tel, T. [1994] "Constrained Frobenius–Perron operator to analyse the dynamics on composed attractors," *Zeitschrift fur Naturforschung A — J. Phys. Sci.* **49**, 1223–1228.