Available online at www.sciencedirect.com
ScienceDirect
Physica D 231 (2007) 1–9
PHYSICA D
www.elsevier.com/locate/physd

A rigorous formalism of information transfer between dynamical system components. I. Discrete mapping

X. San Liang*, Richard Kleeman

Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, NY 10012, USA

Received 24 January 2006; received in revised form 16 October 2006; accepted 23 April 2007
Available online 27 April 2007
Communicated by C.K.R.T. Jones

Abstract

We put the concept of information transfer on a rigorous footing and establish for it a formalism within the framework of discrete maps. The resulting transfer measure possesses a property of directionality or transfer asymmetry as emphasized by Schreiber [T. Schreiber, Measuring information transfer, Phys. Rev. Lett. 85 (2) (2000) 461]; it also verifies the transfer measure for two-dimensional systems, which was obtained by Liang and Kleeman [X.S. Liang, R. Kleeman, Information transfer between dynamical system components, Phys. Rev. Lett. 95 (24) (2005) 244101] through a different avenue. Connections to classical formalisms are explored and applications presented. We find that, in the context of the baker transformation, there is always information flowing from the stretching direction to the folding direction, while no transfer occurs in the opposite direction; we also find that, within the Hénon map system, the transfer from the quadratic component to the linear component is of a simple form as expected on physical grounds. This latter result is unique to our formalism.

Keywords: Information transfer; Frobenius–Perron operator; Baker transformation; Hénon map

1. Introduction

Information transfer between dynamical system components is an important concept for nonlinear multivariate time series coherence analysis (e.g., [2,4,9,20–23]), communication (e.g., [18]), and predictability study [3,11–13] in neurology, atmosphere–ocean science, and many other scientific areas. Information transfer is also seen in our daily life. For example, it has been argued intuitively that, in kneading dough, information is lost continually from the stretching direction to the folding direction (e.g., [14]). A most recent review with particular focus on neurophysiology is Pereda et al. [19]; reviews on a more generic basis of nonlinear signal analysis can be found in the books by Abarbanel [1] and Kantz and Schreiber [10].

Many studies of information transfer are based on mutual information, a natural measure of the independence between two random variables [see [1] and the references there in]. It is well known that mutual information does not possess transfer directionality or asymmetry, a desired property emphasized by Schreiber [22]. A variety of formalisms have bee proposed to address this issue, among which are the time delayed mutual information (e.g., [9,23]) and, in the context of a Markov chain, the more sophisticated transfer entropy [8,22]. These formalisms generally work well in their specific hierarchies and respective contexts.

However, we still lack a formalism to have the unidirectional or asymmetric transfer of information rigorously represented. Here we intend to establish one to fill the gap. We want to show that, when dynamics is fully specified as is the case with many physical problems such as the atmosphere then such a program is feasible. The basic idea has been elucidated within the framework of a two-dimensional (2D) system in [15], which we will refer to as LK05 henceforth. This paper generalizes the 2D formalism to systems of arbitrary dimensionality. As we will soon see, the 2D case is quite special and cannot be directly extended and so a different route will be chosen for the establishment of a general formalism. This paper is concerned with discrete maps only. Systems described by continuous flows

* Corresponding author.
E-mail address: sanliang@courant.nyu.edu (X.S. Liang).

0167-2789/$ - see front matter
doi:10.1016/j.physd.2007.04.002

will be investigated in detail as a limiting case in the second part of this study [16].

In the following we first present the mathematical framework, within which the 2D case studied in LK05 is briefly reviewed, and our strategy for the problem proposed. We then derive the measure of information transfer. Properties are explored (Section 3) to see whether the measure reduces to that of LK05 when dimensionality becomes 2, and whether the transfer thus obtained meets the unidirectionalism requirement. For the sake of comparison, connections are established in Section 4 to the classical formalism. In Section 5, we present two applications with the baker transformation and the Hénon map. We choose these two problems not only because we want to understand how information is transferred between their respective components, but also because they may serve to validate our formalism. This paper is summarized in Section 6.

2. Formalism

2.1. Mathematical framework

Consider a stochastic process $\{\underline{\mathbf{X}}, \tau\}$, with τ positive integers signifying discrete time steps, and $\underline{\mathbf{X}} = (X_1, X_2, \ldots, X_n)$ an array of random variables. We want to know how information is transferred from X_j to X_i, for any $i, j = 1, \ldots, n$, $i \neq j$. Without loss of generality, we need only consider the case $i = 1$, $j = 2$.

Consider an n-dimensional transformation

$$\begin{aligned} &\Phi : \Omega \mapsto \Omega, \\ &(x_1, x_2, \ldots, x_n) \quad \mapsto (\Phi_1(\underline{\mathbf{x}}), \Phi_2(\underline{\mathbf{x}}), \ldots, \Phi_n(\underline{\mathbf{x}})), \end{aligned} \tag{1}$$

where $\underline{\mathbf{x}} = (x_1, \ldots, x_n) \in \Omega$ correspond to the random variables $\underline{\mathbf{X}}$. In this study, the sample space Ω is assumed to be a Cartesian product of $\Omega_1, \Omega_2, \ldots,$ and Ω_n, in which $x_1, x_2, \ldots, x_n$ are respectively lying (Ω_i open in $\mathbb{R}$, $i = 1, \ldots, n$). Let $\rho = \rho(\underline{\mathbf{x}})$ be the joint density of $X_1, X_2, \ldots$, and X_n at step τ. Its evolution is steered by the Frobenius–Perron operator (F–P operator henceforth)

$$\mathcal{P} : L^1(\Omega) \mapsto L^1(\Omega),$$

given by, in a loose sense,

$$\int_\omega \mathcal{P}\rho(\underline{\mathbf{x}}) \mathrm{d}\underline{\mathbf{x}} = \int_{\Phi^{-1}(\omega)} \rho(\underline{\mathbf{x}})\, \mathrm{d}\underline{\mathbf{x}}, \tag{2}$$

for any $\omega \subset \Omega$. [For a rigorous treatment with measure theory, see [14].] When Φ is nonsingular and invertible, Φ can be explicitly written out:

$$\mathcal{P}\rho(\underline{\mathbf{x}}) = \rho\left[\Phi^{-1}(\underline{\mathbf{x}})\right] \left|J^{-1}\right| \tag{3}$$

where J is the determinant of the Jacobian of Φ:

$$J = J(\underline{\mathbf{x}}) = \det\left[\frac{\partial \Phi(x_1, x_2, \ldots, x_n)}{\partial(x_1, x_2, \ldots, x_n)}\right]$$

and J^{-1} its inverse. The joint density $\rho = \rho(\underline{\mathbf{x}})$ defines an entropy (Shannon entropy)

$$H(\underline{\mathbf{X}}) = -\int_\Omega \rho(\underline{\mathbf{x}}) \log \rho(\underline{\mathbf{x}})\, \mathrm{d}\underline{\mathbf{x}}. \tag{4}$$

If a particular component is of interest, say X_1, we need the marginal entropy

$$H_1 = H(\underline{\mathbf{X}}_1) = -\int_{\Omega_1} \rho_1(x_1) \log \rho(x_1)\, \mathrm{d}x_1, \tag{5}$$

where $\rho_1 = \rho_1(x_1) = \int_{\Omega_{2n}} \rho(x_1, x_2, \ldots, x_n) \mathrm{d}x_2 \mathrm{d}x_3 \ldots \mathrm{d}x_n$ is the marginal distribution of X_1. For convenience here we have used the shorthand

$$\Omega_{jn} = \Omega_j \times \Omega_{j+1} \times \cdots \times \Omega_n, \quad 1 \leq j < n,\ (\Omega_{1n} = \Omega),$$

and will keep the convention throughout this paper. The entropy of X_1 evolves as Φ is applied. We are interested in how entropy is transferred from X_2 to X_1 in the course of the evolution. This is the foundation of the previous formalisms of information transfer, which is measured by the amount of entropy thus transferred (cf. Section 4). The objective of this study is, therefore, to find this transfer, which we will denote as $T_{2\to1}$ hereafter, when Φ applies as time goes from step τ to step $\tau + 1$. The following subsections are devoted to the formulation of $T_{2\to1}$.

2.2. The 2D case: A brief summary of the LK05 formalism

In LK05, we have established a formalism for the entropy transfer $T_{2\to1}$ in the context of a 2D system. The fundamental idea is that, the increase in H_1, ΔH_1, can be decomposed into two mutually exclusive parts: one from X_1 itself, written as ΔH_1^*; another one from X_2. Clearly the latter is the transfer $T_{2\to1}$. We may therefore find $T_{2\to1}$ through computing the difference between ΔH_1 and ΔH_1^*. For a system with dynamics given (represented as the transformation Φ), ΔH_1 is fully known. In fact,

$$\begin{aligned} \Delta H_1 &= -\int_{\Omega_1} (\mathcal{P}\rho)_1(x_1)\ \log(\mathcal{P}\rho)_1(x_1)\, \mathrm{d}x_1 - H_1 \\ &= -\int_{\Omega_1} \left(\int_{\Omega_{2n}} \mathcal{P}\rho \mathrm{d}x_2 \ldots \mathrm{d}x_n\right) \\ &\quad \times \log\left(\int_{\Omega_{2n}} \mathcal{P}\rho \mathrm{d}x_2 \ldots \mathrm{d}x_n\right) \mathrm{d}x_1 \\ &\quad + \int_{\Omega_1} \rho_1 \log \rho_1\, \mathrm{d}x_1. \end{aligned} \tag{6}$$

Here we use $(\mathcal{P}\rho)_1$ to indicate the marginal distribution of X_1 at time step $\tau + 1$. If we can find ΔH_1^*, then $T_{2\to1}$ is obtained accordingly.

The evaluation of ΔH_1^* is based on a law governing the evolution of the entropy H of a discrete system under an invertible transformation Φ:

$$\Delta H = E \log |J|, \tag{7}$$

where J is the determinant of the Jacobian of Φ, and $E(\cdot)$ the mathematical expectation over the whole sample space. Eq. (7) was obtained by LK05. It states that the entropy increase for a discrete system upon one application of an invertible transformation is simply the average logarithm of the rate of area change under that transformation. Since (7) applies to any systems with invertible transformations, one may argue heuristically that, when only X_1 is considered, the change of the marginal entropy H_1, or ΔH_1^* as we previously denoted, is

$$\Delta H_1^* = E \log |J_1|, \tag{8}$$

where $J_1 = \frac{\partial \Phi_1(\mathbf{\underline{x}})}{\partial x_1}$, and $E(\cdot)$ is, as in (7), over the whole space Ω. Subtracting (8) from (6), we arrive at

$$T_{2\to 1} = \Delta H_1 - \Delta H_1^* = \Delta H_1 - E \log |J_1|. \tag{9}$$

This is the entropy transfer from X_2 to X_1, holding when $n = 2$ and Φ_1 is invertible.

2.3. The general case

The formalism (9) of entropy transfer seems to be natural. Two limitations, however, prevent it from being applicable to general problems. The first regards the invertibility of Φ_1. In many interesting problems, even though the mapping Φ as a whole is invertible, its components, say Φ_1, are more often than not noninvertible. The two applications which we will present toward the end of this paper are just such examples. Eq. (9) must be extended to have noninvertibility included.

A bigger problem with the above formalism lies in the fact that the strategy to obtain (9) only works for 2D systems. When the dimensionality $n \geq 3$, it becomes invalid, otherwise the transfer thus obtained would be the bulk transfer to X_1 from all other components. We would not get $T_{2\to 1}$ unless we can differentiate the contributions of X_j, $j = 3, 4, \ldots n$, from that of X_2, which is, unfortunately, by no means an easy task.

Out of the problem comes the solution. The ΔH_1^* in (9), on the other hand, may be equally understood as the evolution of H_1 with the influence from X_2 excluded, namely, the change of H_1 with x_2 frozen at step τ. In this spirit, we may then partition the mechanisms governing the evolution of H_1 into two disjoint subsets: One is the transfer from X_2, another one the evolution without influence from X_2. The transfer is then the difference between the two. This partitioning does not have any restraints on n. The formalism of transfer is therefore applicable to systems of arbitrary dimensionality.

Given the dynamics, it is easy to obtain ΔH_1. The key to the formalism is thus to find the contribution to the increase of H_1 with X_2 excluded. In formal language, this is the entropy increase in direction 1 as the system goes from step τ to $\tau + 1$ under Φ with x_2 frozen instantaneously at time step τ, given $X_1(\tau)$. We denote it as $\Delta H_{1\not{2}}$. For later convenience, this index notation is extended to any symbol in the form $\not{j}$ to signify that component j is frozen, or that component j is excluded from a set of n independent variables. When ρ is referred to, this simply means marginalization of that component, e.g.,

$$\rho_{\not{2}} = \rho_{\not{2}}(x_1, x_3, \ldots, x_n) = \int_{\Omega_2} \rho(x_1, x_2, x_3, \ldots, x_n) \mathrm{d}x_2, \tag{10}$$

$$\rho_{\not{1}\not{2}} = \rho_{\not{1}\not{2}}(x_3, \ldots, x_n) = \int_{\Omega_1 \times \Omega_2} \rho(x_1, x_2, x_3, \ldots, x_n) \, \mathrm{d}x_1 \mathrm{d}x_2. \tag{11}$$

We will use this convention henceforth without further clarification.

To find $\Delta H_{1\not{2}}$, we need to evaluate $H_{1\not{2}}(\tau + 1)$, the marginal entropy for the first component evolved from H_1 with contribution from X_2 excluded from step τ to $\tau + 1$. For this purpose, consider the quantity

$$f \equiv -\log(\mathcal{P}\rho)_{1\not{2}}(y_1), \tag{12}$$

where $y_1 = \Phi_1(\mathbf{\underline{x}})$, and $(\mathcal{P}\rho)_{1\not{2}}(y_1)$ is the marginal density in direction 1 at step $\tau + 1$, as the density $\rho_{\not{2}}$ evolves from step τ to step $\tau + 1$ under the transformation:

$$\Phi_{\not{2}} : \begin{cases} y_1 = & \Phi_1(x_1, x_2, x_3, \ldots, x_n) \\ y_3 = & \Phi_3(x_1, x_2, x_3, \ldots, x_n) \\ \vdots & \vdots \\ y_n = & \Phi_n(x_1, x_2, x_3, \ldots, x_n) \end{cases} \tag{13}$$

i.e., the map Φ with x_2 frozen instantaneously at τ as a parameter. Note here we use $y_1 = \Phi_1(\mathbf{\underline{x}})$ to represent the state of component 1 at step $\tau + 1$ (x_1 is for that at step τ). We do not use x_1 with some superscript or subscript in order to avoid any possible confusion in distinguishing the states of X_1 at these two time steps.

Following the definition of Shannon entropy, $H_{1\not{2}}(\tau + 1)$ is in the form of some average (expectation) of f. In other words, it is equal to the integration of f times some probability density function over the corresponding sample space. The first density to be multiplied is $(\mathcal{P}\rho)_{1\not{2}}(y_1)$, but f also depends on x_2 as well as y_1. We need another density for X_2. Recall that the freezing of x_2 is performed on interval $[\tau, \ \tau + 1]$, given all other components at time step τ. What we need is therefore the conditional density of X_2 on $X_1, X_3, \ldots, X_n$ at τ, $\rho(x_2|x_1, x_3, \ldots, x_n)$. Among the newly introduced variables, $x_3, x_4, \ldots, x_n$ should be averaged out, so another density $\rho_{3\ldots n}(x_3, \ldots, x_n)$ should be multiplied, while by (13) x_1 can be viewed as a function of y_1 and $x_3, x_4, \ldots, x_n$. Based on these arguments, the marginal entropy for the first component evolved from H_1 with contribution from X_2 excluded from step τ to $\tau + 1$ is

$$H_{1\not{2}}(\tau + 1) = -\int_\Omega (\mathcal{P}\rho)_{1\not{2}}(y_1) \cdot \log(\mathcal{P}\rho)_{1\not{2}}(y_1) \cdot \rho(x_2|x_1, x_3, \ldots, x_n) \cdot \rho_{3\ldots n}(x_3, \ldots, x_n) \times \mathrm{d}y_1 \mathrm{d}x_2 \mathrm{d}x_3 \ldots \mathrm{d}x_n. \tag{14}$$

So

$$\Delta H_{1\not{2}} = H_{1\not{2}}(\tau + 1) - H_1(\tau). \tag{15}$$

The information transfer from X_2 to X_1 is the difference between ΔH_1 and $\Delta H_{1\not{2}}$. Subtraction of (15) from (6) yields

$$\begin{aligned} T_{2\to 1} = & -\int_{\Omega_1} (\mathcal{P}\rho)_1(y_1) \cdot \log(\mathcal{P}\rho)_1(y_1) \, \mathrm{d}y_1 \\ & + \int_\Omega (\mathcal{P}\rho)_{1\not{2}}(y_1) \cdot \log(\mathcal{P}\rho)_{1\not{2}}(y_1) \\ & \cdot \rho(x_2|x_1, x_3, \ldots, x_n) \\ & \cdot \rho_{3\ldots n}(x_3, \ldots, x_n) \, \mathrm{d}y_1 \mathrm{d}x_2 \mathrm{d}x_3 \ldots \mathrm{d}x_n. \end{aligned} \tag{16}$$

Note that $y_1 = \Phi_1(\mathbf{\underline{x}})$ and x_1 are the state variables at different time steps, and that the conditional probability of

x_2 is on x_1 instead of y_1. In arriving at (16), no issue about the invertibility of Φ or any of its components has been invoked. The formalism is therefore valid for a transformation of arbitrary dimensionality.

Likewise, for any $i, j = 1, 2, \ldots, n$, $i \neq j$, the entropy transfer from X_j to X_i is

$$\begin{aligned} T_{j\to i} = &-\int_{\Omega_i} (\mathcal{P}\rho)_i(y_i)\cdot\log(\mathcal{P}\rho)_i(y_i)\,\mathrm{d}y_i \\ &+\int_{\Omega} (\mathcal{P}\rho)_{i\not j}(y_i)\cdot\log(\mathcal{P}\rho)_{1\not j}(y_i) \\ &\cdot\rho(x_j|x_1,x_2,\ldots,x_{j-1},x_{j+1},\ldots,x_n) \\ &\cdot\rho_{\not i\not j}\,\mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_{i-1}\mathrm{d}y_i\mathrm{d}x_{i+1}\ldots\mathrm{d}x_n. \end{aligned} \tag{17}$$

3. Properties

The transfer measure developed above possesses some important properties. This section presents two of them in the form of theorems. Again, in the interest of the readership, we will avoid considering measurability whenever raised in the derivations or proofs.

First, we expect our formalism verifies the case for 2D systems, which we obtained rigorously in LK05 via a different route. This forms the following theorem:

Theorem 1. *When $n = 2$ and Φ_1 is invertible, $\Delta H_{1\not 2} = E\log|J_1|$.*

Proof. If $n = 2$,

$$\begin{aligned} \Delta H_{1\not 2} = &-\iint_{\Omega_1\times\Omega_2} (\mathcal{P}\rho)_{1\not 2}(y_1) \\ &\cdot\log(\mathcal{P}\rho)_{1\not 2}(y_1)\cdot\rho(x_2|x_1)\,\mathrm{d}y_1\mathrm{d}x_2 \\ &+\int_{\Omega_1}\rho_1\log\rho_1\mathrm{d}x_1, \end{aligned}$$

where $y_1 = \Phi_1(x_1, x_2)$, and $(\mathcal{P}\rho)_{1\not 2}$ the marginal distribution of X_1 evolving from $\rho_{\not 2} = \rho_1$ upon one transformation of $\Phi_{\not 2} = \Phi_1$. When Φ_1 is invertible, $J_1 = \frac{\partial\Phi_1}{\partial x_1} \neq 0$, by (3),

$$\begin{aligned} (\mathcal{P}\rho)_{1\not 2}(y_1) &= \rho\left[\Phi_1^{-1}(y_1,x_2)\right]\left|J_1^{-1}\right| \\ &= \rho_1(x_1)\left|J_1^{-1}\right|. \end{aligned} \tag{18}$$

So

$$\begin{aligned} \Delta H_{1\not 2} &= -\iint \rho_1(x_1)\left|J_1^{-1}\right|\log\left(\rho_1(x_1)\left|J_1^{-1}\right|\right) \\ &\quad\rho(x_2|x_1)\,|J_1|\,\mathrm{d}x_1\mathrm{d}x_2 + \int\rho_1\log\rho_1\,\mathrm{d}x_1 \\ &= -\iint\rho_1(x_1)\,\rho(x_2|x_1)\,\log\left|J_1^{-1}\right|\,\mathrm{d}x_1\mathrm{d}x_2 \\ &= \iint\rho(x_1,x_2)\log|J_1|\,\mathrm{d}x_1\mathrm{d}x_2 \\ &= E\log|J_1|.\quad\square \end{aligned} \tag{19}$$

Information transfer is by observation not symmetric between two components, and Schreiber [22] emphasized that a faithful formalism must be able to reflect this asymmetry. In this context, Schreiber's argument can be put in a more concrete and quantitative way:

Theorem 2. *If Φ_1 is independent of x_2, then $T_{2\to 1} = 0$; in the same time $T_{1\to 2}$ need not be zero unless Φ_2 does not rely on x_1.*

Proof. What we need to show is that

$$H_1(\tau+1) = -\int_{\Omega_1}(\mathcal{P}\rho)_1(y_1)\cdot\log(\mathcal{P}\rho)_1(y_1)\,\mathrm{d}y_1 \tag{20}$$

$$\begin{aligned} H_{1\not 2}(\tau+1) = &-\int_{\Omega}(\mathcal{P}\rho)_{1\not 2}(y_1)\cdot\log(\mathcal{P}\rho)_{1\not 2}(y_1) \\ &\cdot\rho(x_2|x_1,x_3,\ldots,x_n) \\ &\cdot\rho_{3\ldots n}(x_3,\ldots,x_n)\,\mathrm{d}y_1\mathrm{d}x_2\mathrm{d}x_3\ldots\mathrm{d}x_n \end{aligned} \tag{21}$$

are identical when Φ_1 is independent of x_2. For an arbitrary subset of Ω_1, $\omega_1 \subset \Omega_1$, we have

$$\begin{aligned} \int_{\omega_1}(\mathcal{P}\rho)_{1\not 2}(x_1)\,\mathrm{d}x_1 &= \int_{\omega_1\times\Omega_{3n}}(\mathcal{P}\rho)_{\not 2}(x_1,x_3,\ldots,x_n) \\ &\quad\times\mathrm{d}x_1\mathrm{d}x_3\ldots\mathrm{d}x_n \\ &= \int_{\Phi_{\not 2}^{-1}(\omega_1\times\Omega_{3n})}\rho_{\not 2}(x_1,x_3,\ldots,x_n) \\ &\quad\times\mathrm{d}x_1\mathrm{d}x_3\ldots\mathrm{d}x_n \end{aligned}$$

by (2) and the definition of $(\mathcal{P}\rho)_{\not 2}$. Note[1]

$$\Phi_{\not 2}^{-1}(\omega_1\times\Omega_{3n}) = \Phi_{1\not 2}^{-1}\omega_1\times\Omega_{3n}.$$

So

$$\begin{aligned} \int_{\omega_1}(\mathcal{P}\rho)_{1\not 2}(x_1)\,\mathrm{d}x_1 &= \int_{\Phi_{1\not 2}^{-1}\omega_1}\mathrm{d}x_1\int_{\Omega_{3n}}\rho_{\not 2}(x_1,x_3,\ldots,x_n) \\ &\quad\times\mathrm{d}x_3\ldots\mathrm{d}x_n \\ &= \int_{\Phi_{1\not 2}^{-1}\omega_1}\rho_1(x_1)\,\mathrm{d}x_1 \\ &= \int_{\Phi_1^{-1}\omega_1}\rho_1(x_1)\,\mathrm{d}x_1 \end{aligned}$$

because Φ_1 (and hence Φ_1^{-1}) has no dependence on x_2.

On the other hand,

$$\begin{aligned} \int_{\omega_1}(\mathcal{P}\rho)_1(x_1)\,\mathrm{d}x_1 &= \int_{\omega_1\times\Omega_{2n}}\mathcal{P}\rho(\underline{\mathbf{x}})\,\mathrm{d}\underline{\mathbf{x}} \\ &= \int_{\Phi^{-1}(\omega_1\times\Omega_{2n})}\rho(\underline{\mathbf{x}})\,\mathrm{d}\underline{\mathbf{x}} \\ &= \int_{\Phi_1^{-1}\omega_1\times\Omega_{2n}}\rho(\underline{\mathbf{x}})\,\mathrm{d}\underline{\mathbf{x}} \\ &= \int_{\Phi_1^{-1}\omega_1}\mathrm{d}x_1\int_{\Omega_{2n}}\rho(\underline{\mathbf{x}})\mathrm{d}x_2\mathrm{d}x_3\ldots\mathrm{d}x_n \\ &= \int_{\Phi_1^{-1}\omega_1}\rho_1(x_1)\,\mathrm{d}x_1. \end{aligned}$$

So

$$\int_{\omega_1}(\mathcal{P}\rho)_{1\not 2}(x_1)\,\mathrm{d}x_1 = \int_{\omega_1}(\mathcal{P}\rho)_1(x_1)\,\mathrm{d}x_1$$

[1] Throughout this proof, $\Phi_{1\not 2}^{-1}\omega_1$ and $\Phi_1^{-1}\omega_1$ should be understood respectively as their first factors (i.e., the projection of the Cartesian product on the subspace of x_1). We abuse the notation for simplicity.

for any $\omega_1 \subset \Omega_1$. This is to say,

$$(\mathcal{P}\rho)_{1\not{2}} = (\mathcal{P}\rho)_1 \tag{22}$$

almost everywhere when Φ_1 is independent of x_2, i.e., the case of inequality has measure zero. (A rigorous treatment requires a consideration of measurability, which is beyond the scope of this study.) Eq. (21) thus becomes

$$\begin{aligned} H_{1\not{2}}(\tau+1) = &-\int_\Omega (\mathcal{P}\rho)_1(y_1)\cdot\log(\mathcal{P}\rho)_1(y_1)\\ &\cdot\rho(x_2|x_1,x_3,\ldots,x_n)\\ &\cdot\rho_{3\ldots n}(x_3,\ldots,x_n)\,\mathrm{d}y_1\mathrm{d}x_2\mathrm{d}x_3\ldots\mathrm{d}x_n. \end{aligned}$$

Observe that $(\mathcal{P}\rho)_1$ has no dependence on x_i, for $i = 2,\ldots,n$. So

$$\begin{aligned} H_{1\not{2}}(\tau+1) &= -\int_{\Omega_1} (\mathcal{P}\rho)_1(y_1)\cdot\log(\mathcal{P}\rho)_1(y_1)\,\mathrm{d}y_1\\ &\quad\cdot\left[\int_{\Omega_2}\int_{\Omega_{3n}} \rho(x_2|x_1,x_3,\ldots,x_n)\right.\\ &\quad\left.\cdot\rho_{3\ldots n}(x_3,\ldots,x_n)\mathrm{d}x_2\mathrm{d}x_3\ldots\mathrm{d}x_n\right]\\ &= -\int_{\Omega_1} (\mathcal{P}\rho)_1(y_1)\cdot\log(\mathcal{P}\rho)_1(y_1)\,\mathrm{d}y_1\\ &= H_1(\tau+1), \end{aligned}$$

where we have used the fact that the part in the square bracket integrates to 1, as x_2 can be first integrated out

$$\int_{\Omega_2} \rho(x_2|x_1,x_3,\ldots,x_n)\mathrm{d}x_2 = 1,$$

and the remaining integral $\int_{\Omega_{3n}} \rho_{3\ldots n}(x_3,\ldots,x_n)\mathrm{d}x_3\ldots\mathrm{d}x_n$ is also equal to 1. Therefore

$$T_{2\to1} = [H_1(\tau+1) - H_1(\tau)] - [H_{1\not{2}}(\tau+1) - H_1(\tau)] = 0,$$

which is what we want to prove. □

Note that the above theorem holds not just between X_1 and X_2, but between X_i and X_j for any $i \neq j$.

Corollary 3. *If Φ_i is independent of x_j, $1 \le i, j \le n$, $i \neq j$, then $T_{j\to i} = 0$.*

To prove, we just need reorder the components of $\underline{\mathbf{X}}$ to have X_i and X_j placed in the first and the second slots of the array so that the above proof applies.

Theorem 2 and Corollary 3 state that transfer of information in one direction yields no hint about the other direction. Particularly, when X_i evolves independently of X_j, there is no transfer from X_j to X_i, but at the same time there could be information flowing in the opposite direction provided that X_j relies on X_i to grow. This property of transfer asymmetry or unidirectionalism makes information transfer conceptually distinct from transfers of other physical quantities such as energy, in which transfer symmetry holds.

4. Connection to previous formalisms

We have adopted a route distinctly different from the classical ones in deriving information transfer. However, it is interesting to note that our formalism is physically consistent with the classical formalism. Particularly, it is consistent with Schreiber's transfer entropy [22]. We have mentioned this in LK05. In the following we show how.

Let P denote the probability mass function. In the case of a Markov chain of order one, the transfer entropy from X_2 to X_1 is, at time step τ,

$$T^S_{2\to1} = \sum P(x_1^{\tau+1}, x_1^\tau, x_2^\tau)\ \log\frac{P(x_1^{\tau+1}|x_1^\tau, x_2^\tau)}{P(x_1^{\tau+1}|x_1^\tau)}. \tag{23}$$

Here the symbol T corresponds to our transfer, and the superscript S is utilized to signify Schreiber's formalism. $T^S_{2\to1}$ is a relative entropy-like quantity (see, for example, [5]) which characterizes the incorrectness when the probability of X_1 at time step τ conditioned on the measurements at previous time steps is taken as the probability of X_1 given the measurements of both X_1 and X_2 at their previous time steps. Schreiber uses it to measure the information transfer from X_2 to X_1.

Notice the transfer entropy (23) may also be written in the form of a difference, as in (16),

$$T^S_{2\to1} = \Delta H^S_1 - \Delta H^S_{1|2}, \tag{24}$$

where

$$\begin{aligned} \Delta H^S_1 &= -\sum P(x_1^{\tau+1}, x_1^\tau)\log P(x_1^{\tau+1}, x_1^\tau)\\ &\quad + \sum P(x_1^{\tau+1}, x_1^\tau)\log P(x_1^\tau), \end{aligned} \tag{25}$$

and

$$\begin{aligned} \Delta H^S_{1|2} &= -\sum P(x_1^{\tau+1}, x_1^\tau, x_2^\tau)\log P(x_1^{\tau+1}, x_1^\tau, x_2^\tau)\\ &\quad + \sum P(x_1^{\tau+1}, x_1^\tau, x_2^\tau)\log P(x_1^\tau, x_2^\tau). \end{aligned} \tag{26}$$

For ΔH^S_1, the second term on the right hand side is

$$\begin{aligned} -\sum P(x_1^{\tau+1}, x_1^\tau)\log P(x_1^\tau) &= -\sum P(x_1^\tau)\log P(x_1^\tau)\\ &= H_1(\tau); \end{aligned}$$

the first term is also like entropy, but at a time step between τ and $\tau+1$. We denote it $H_1(\tau+\frac{1}{2})$ for the time being. So

$$\Delta H^S_1 = H_1\left(\tau+\frac{1}{2}\right) - H_1(\tau) \tag{27}$$

is a kind of entropy increase. To see the physical meaning of $\Delta H^S_{1|2}$, introduce two quantities

$$\begin{aligned} A &= \sum P(x_1^{\tau+1}, x_1^\tau, x_2^\tau)\log P(x_2^\tau),\\ B &= \sum P(x_1^\tau, x_2^\tau)\log P(x_2^\tau). \end{aligned}$$

It is easy to show that both A and B are equal to $-H_2(\tau)$. We may then have (26) first plus A then minus B to get

$$\begin{aligned} \Delta H^S_{1|2} &= \left[-\sum P(x_1^{\tau+1}, x_1^\tau, x_2^\tau)\log P(x_1^{\tau+1}, x_1^\tau, x_2^\tau) + A\right]\\ &\quad - \left[-\sum P(x_1^\tau, x_2^\tau)\log P(x_1^\tau, x_2^\tau) + B\right]\\ &= -\sum P(x_1^{\tau+1}, x_1^\tau, x_2^\tau)\log P(x_1^{\tau+1}, x_1^\tau|x_2^\tau)\\ &\quad + \sum P(x_1^\tau, x_2^\tau)\log P(x_1^\tau|x_2^\tau). \end{aligned}$$

It is now clear that the last term on the right hand side is the conditional entropy of X_1 on X_2 at time step τ, $H_{1|2}(\tau)$, while the first term may also be understood as the conditional entropy of X_1 on X_2 at some time step between τ and $\tau+1$, denoted $H_{1|2}(\tau+\frac{1}{2})$. So $\Delta H^S_{1|2}$ describes some entropy increment of X_1 conditioned on X_2.

Schreiber's transfer entropy is the difference between ΔH^S_1 and $\Delta H^S_{1|2}$. These two terms correspond to our ΔH_1 in (6) and $\Delta H_{1\not{2}}$ in (15), respectively. In a sense, the freezing of x_2 instantaneously as time goes from τ to $\tau+1$ can be viewed as a kind of conditioning on X_2. Our formalism is thus physically consistent with the transfer entropy (23) or (24) in the context of a Markov chain of order one.

But our formalism is obviously different from the transfer entropy. The difference lies in the following two aspects: (a) The entropy increases in (24) are those from step τ to somewhere between τ and $\tau+1$, while in our formalism, both ΔH_1 and $\Delta H_{1\not{2}}$ describe the entropy variations from step τ to step $\tau+1$; (b) although qualitatively there is some similarity between the conditioning in $\Delta H^S_{1|2}$ and the freezing in $\Delta H_{1\not{2}}$, quantitatively they are not identical. These differences may give different results for the same problem. For example, there is no property like Theorem 1 with (24). We will see more examples in the following applications.

5. Applications

We now calculate (16) or (17) for the well known baker transformation and Hénon map. We choose these two problems not only because their information transfers per se are interesting, but also because they may in some sense serve to validate our formalism. Both maps have one or more noninvertible components, and so the simple formalism (9) does not apply.

5.1. Baker transformation

It has been argued intuitively that, in applying the baker transformation which mimics the kneading of dough, information flows continuingly from the stretching direction to the folding direction, while no transfer occurs in the opposite direction (see [14]; similar arguments can also be seen in [6,7,17]). We mentioned this in the beginning of the paper, and want to see whether this is the case with our formalism.

The baker transformation is defined as a mapping $\Phi : \Omega \to \Omega$, $\Omega = [0,1] \times [0,1]$, given by

$$\Phi(x_1,x_2) = \begin{cases} \left(2x_1,\ \dfrac{x_2}{2}\right) & 0 \le x_1 \le \dfrac{1}{2},\ 0 \le x_2 \le 1 \\ \left(2x_1 - 1,\ \dfrac{1}{2}x_2 + \dfrac{1}{2}\right) & \dfrac{1}{2} < x_1 \le 1,\ 0 \le x_2 \le 1. \end{cases} \tag{28}$$

It is invertible, and measure preserving ($J = 1$), as is easy to check. By Eq. (7) this means that its entropy stays unchanged. (But one of its components is not. See below.) To compute the information transfer, we need the F–P operator $\mathcal{P}$, which can easily evaluated by taking advantage of the invertibility (cf. Appendix):

$$\mathcal{P}(x_1,x_2) = \begin{cases} \rho\left(\dfrac{x_1}{2}, 2x_2\right), & 0 \le x_2 < \dfrac{1}{2}, \\ \rho\left(\dfrac{1+x_1}{2}, 2x_2 - 1\right), & \dfrac{1}{2} \le x_2 \le 1. \end{cases} \tag{29}$$

Consider the transfer from X_2 to X_1 first. Upon one transformation, the marginal density increases from

$$\rho_1 = \int_0^1 \rho(x_1,x_2)\mathrm{d}x_2$$

to

$$\begin{aligned} \int_0^1 \mathcal{P}\rho(x_1,x_2)\mathrm{d}x_2 &= \int_0^{1/2} \rho\left(\frac{x_1}{2}, 2x_2\right)\mathrm{d}x_2 \\ &\quad + \int_{1/2}^1 \rho\left(\frac{x_1+1}{2}, 2x_2 - 1\right)\mathrm{d}x_2 \\ &= \frac{1}{2}\int_0^1 \left[\rho\left(\frac{x_1}{2}, x_2\right) + \rho\left(\frac{x_1+1}{2}, x_2\right)\right]\mathrm{d}x_2 \\ &= \frac{1}{2}\left[\rho_1\left(\frac{x_1}{2}\right) + \rho_1\left(\frac{x_1+1}{2}\right)\right], \end{aligned} \tag{30}$$

where (29) has been used in the derivation.

The baker transformation as a whole is invertible. Its x_1 direction, however, is not. Consider X_1 only, the transformation reduces to a dyadic mapping, $\Phi_1 : [0,1] \to [0,1]$, $\Phi_1(x_1) = 2x_1 \pmod 1$. It has an F–P operator

$$(\mathcal{P}\rho)_{1\not{2}}(x_1) = \frac{1}{2}\left[\rho_1\left(\frac{x_1}{2}\right) + \rho_1\left(\frac{1+x_1}{2}\right)\right] \tag{31}$$

(see Appendix for details). This is exactly the same as (30), implying that

$$T_{2\to1} = 0, \tag{32}$$

which is just as expected.

Now compute the transfer from X_1 to X_2. As above, first compute the marginal distribution

$$\int_0^1 \mathcal{P}\rho(x_1,x_2)\,\mathrm{d}x_1 = \begin{cases} \displaystyle\int_0^1 \rho\left(\frac{x_1}{2}, 2x_2\right)\,\mathrm{d}x_1, & 0 \le x_2 < \dfrac{1}{2}; \\ \displaystyle\int_0^1 \rho\left(\frac{x_1+1}{2}, 2x_2 - 1\right)\,\mathrm{d}x_1, & \dfrac{1}{2} \le x_2 \le 1. \end{cases} \tag{33}$$

This substituted in

$$\begin{aligned} \Delta H_2 = &-\int_0^1\int_0^1 \mathcal{P}\rho(x_1,x_2) \\ &\cdot\left[\log\left(\int_0^1 \mathcal{P}\rho(\lambda,x_2)\mathrm{d}\lambda\right)\right]\mathrm{d}x_1\mathrm{d}x_2 \\ &+ \int_0^1\int_0^1 \rho(x_1,x_2) \\ &\cdot\left[\log\left(\int_0^1 \rho(\lambda,x_2)\mathrm{d}\lambda\right)\right]\mathrm{d}x_1\mathrm{d}x_2, \end{aligned} \tag{34}$$

after a series of transformation of variables, gives

$$\Delta H_2 = -\log 2 + (I + II), \tag{35}$$

where

$$I = \int_0^1 \int_0^{1/2} \rho(x_1, x_2) \cdot \left[\log \frac{\int_0^1 \rho(\lambda, x_2)\mathrm{d}\lambda}{\int_0^{1/2} \rho(\lambda, x_2)\mathrm{d}\lambda} \right] \mathrm{d}x_1 \mathrm{d}x_2, \tag{36}$$

$$II = \int_0^1 \int_{1/2}^1 \rho(x_1, x_2) \cdot \left[\log \frac{\int_0^1 \rho(\lambda, x_2)\mathrm{d}\lambda}{\int_{1/2}^1 \rho(\lambda, x_2)\mathrm{d}\lambda} \right] \mathrm{d}x_1 \mathrm{d}x_2. \tag{37}$$

Note both I and II are nonnegative, because $\rho(x_1, x_2) \geq 0$ and

$$\int_0^1 \rho(x_1, x_2)\, \mathrm{d}x_1 \geq \int_0^{1/2} \rho(x_1, x_2)\, \mathrm{d}x_1 \tag{38}$$

$$\int_0^1 \rho(x_1, x_2)\, \mathrm{d}x_1 \geq \int_0^{1/2} \rho(x_1, x_2)\, \mathrm{d}x_1. \tag{39}$$

Moreover, these equalities cannot hold simultaneously. So $I+II$ is strictly positive.

On the other hand, in the x_2 direction the transformation is invertible and J_2 is equal to a constant $\frac{1}{2}$. By Theorem 1,

$$\Delta H_{2\backslash\!\!\!1} = E \log \frac{1}{2} = -\log 2. \tag{40}$$

So,

$$T_{1\to 2} = \Delta H_2 - \Delta H_{2\backslash\!\!\!1} = I + II > 0, \tag{41}$$

i.e., there is always information flowing from X_1 to X_2.

Eqs. (32) and (41) show that information flows continuously from the stretching direction to the folding direction ($T_{1\to 2} > 0$), while no transfer occurs in the opposite direction ($T_{2\to 1} = 0$). These are just what we have expected with the baker transformation. Our formalism thus produces a result agreeing well with intuition.

5.2. *Information transfer in the Hénon map*

The Hénon map $\Phi = (\Phi_1, \Phi_2) : \mathbb{R}^2 \mapsto \mathbb{R}^2$ is defined such that

$$\begin{cases} \Phi_1(x_1, x_2) = 1 + x_2 - a x_1^2, \\ \Phi_2(x_1, x_2) = b x_1, \end{cases} \tag{42}$$

with $a > 0$, $b > 0$. Like the baker transformation, it is also invertible. In fact, the inverse mapping is

$$\Phi^{-1}(x_1, x_2) = \left(\frac{x_2}{b},\ x_1 - 1 + \frac{a}{b^2} x_2^2 \right). \tag{43}$$

By (3),

$$\begin{aligned} \mathcal{P}\rho(x_1, x_2) &= \rho(\Phi^{-1}(x_1, x_2)) |J^{-1}| \\ &= \frac{1}{b} \rho \left(\frac{x_2}{b},\ x_1 - 1 + \frac{a}{b^2} x_2^2 \right) \\ &\equiv \frac{1}{b} \rho(y,\ x_1 - 1 + a y^2), \end{aligned} \tag{44}$$

where for convenience y is used to denote $\frac{x_2}{b}$. We hereafter compute the transfers between X_1 and X_2. (a) $\mathrm{T}_{2\to 1}$: *Transfer from the linear component to the quadratic component*

According to (16), we need to evaluate $(\mathcal{P}\rho)_1$ and $(\mathcal{P}\rho)_{1\backslash\!\!\!2}$. The former is

$$\begin{aligned} (\mathcal{P}\rho)_1(x_1) &= \int_{\mathbb{R}} \mathcal{P}\rho(x_1, x_2) \mathrm{d}x_2 \\ &= \int_{\mathbb{R}} \frac{1}{b} (y, x_1 - 1 + a y^2) \mathrm{d}x_2 \\ &= \int_{\mathbb{R}} \rho(y, x_1 - 1 + a y^2) \mathrm{d}y \\ &= \rho_2(x_1). \end{aligned}$$

(Note the argument: it is x_1, not x_2.) To compute $(\mathcal{P}\rho)_{1\backslash\!\!\!2}$, let

$$x_1' \equiv \Phi_1(x_1) = 1 + x_2 - a x_1^2,$$

where x_2 appears as a parameter. From this it is easy to obtain

$$\Phi_1^{-1}((-\infty, x_1']) = \left(-\infty, -\sqrt{\frac{1 + x_2 - x_1'}{a}} \right] \cup \left[\sqrt{\frac{1 + x_2 - x_1'}{a}}, \infty \right).$$

By the formula for F–P operator evaluation (see [14])

$$\begin{aligned} (\mathcal{P}\rho)_{1\backslash\!\!\!2}(x_1') &= \frac{\mathrm{d}}{\mathrm{d}x_1'} \int_{\Phi_1^{-1}((-\infty, x_1'])} \rho_1(s) \mathrm{d}s \\ &= \frac{\mathrm{d}}{\mathrm{d}x_1'} \int_{-\infty}^{-\sqrt{\frac{1+x_2-x_1'}{a}}} \rho_1(s) \mathrm{d}s + \frac{\mathrm{d}}{\mathrm{d}x_1'} \int_{\sqrt{\frac{1+x_2-x_1'}{a}}}^{\infty} \rho_1(s) \mathrm{d}s \\ &= \frac{1}{2\sqrt{a(1 + x_2 - x_1')}} [\rho]_1 \left(-\sqrt{\frac{1 + x_2 - x_1'}{a}} \right) + \rho_1 \left(\sqrt{\frac{1 + x_2 - x_1'}{a}} \right) \quad (x_1' < 1 + x_2) \\ &= \frac{1}{2a|x_1|} [\rho_1(-x_1) + \rho_1(x_1)]. \\ &\quad (\text{recall } x_1' = 1 + x_2 - a x_1^2) \end{aligned}$$

This and the above expression for $(\mathcal{P}\rho)_1$, together with the fact $J_1 = \frac{\partial \Phi_1}{\partial x_1} = -2 a x_1$, substituted into (16) gives

$$\begin{aligned} T_{2\to 1} = &-\int_{\mathbb{R}} (\mathcal{P}\rho)_1(x_1) \cdot \log(\mathcal{P}\rho)_1(x_1)\, \mathrm{d}x_1 \\ &+ \iint_{\mathbb{R}\times\mathbb{R}} (\mathcal{P}\rho)_{1\backslash\!\!\!2}(x_1') \log(\mathcal{P}\rho)_{1\backslash\!\!\!2}(x_1') \\ &\cdot \rho(x_2|x_1) \cdot |J_1|\, \mathrm{d}x_1 \mathrm{d}x_2 \end{aligned}$$

$$= -\int_{\mathbb{R}} \rho_2(x_1)\log\rho_2(x_1)\,\mathrm{d}x_1 + \int_{\mathbb{R}} \frac{\rho_1(-x_1)+\rho_1(x_1)}{2a|x_1|} \cdot \log\frac{\rho_1(-x_1)+\rho_1(x_1)}{2a|x_1|} \cdot |-2ax_1|\,\mathrm{d}x_1$$
$$= H_2 + 2\int_0^\infty [\rho_1(-x_1)+\rho_1(x_1)] \cdot \log\frac{\rho_1(-x_1)+\rho_1(x_1)}{2ax_1}\,\mathrm{d}x_1. \tag{45}$$

Eq. (45) states that the entropy transferred from the linear component (X_2) to the quadratic component (X_1) is equal to the entropy of X_2 (H_2) modified by a part arising from the involvement of x_1 in Φ_1 [cf. (42)].

(b) $T_{1\to2}$: *Transfer from the quadratic component to the linear component*

From (17),

$$T_{1\to2} = -\int_{\mathbb{R}} (\mathcal{P}\rho)_2(x_2)\cdot\log(\mathcal{P}\rho)_2(x_2)\mathrm{d}x_2 \int_{\mathbb{R}} (\mathcal{P}\rho)_{2\backslash}(x_2')\cdot\log(\mathcal{P}\rho)_{2\backslash}(x_2') \cdot\rho(x_1|x_2)\cdot|J_2|\ \mathrm{d}x_1\mathrm{d}x_2,$$

where $x_2' = \Phi_2(x_1, x_2)$. Notice $J_2 = \frac{\partial\Phi_2}{\partial x_2} = 0$, and

$$(\mathcal{P}\rho)_2(x_2) = \int_{\mathbb{R}} \mathcal{P}\rho(x_1,x_2)\,\mathrm{d}x_1 = \int_{\mathbb{R}} \frac{1}{b}\rho\left(\frac{x_2}{b}, x_1 - 1 + a\frac{x_2^2}{b^2}\right)\mathrm{d}x_1 = \frac{1}{b}\int_{\mathbb{R}}\rho(y,\xi)\,\mathrm{d}\xi = \frac{1}{b}\rho_1\left(\frac{x_2}{b}\right).$$

So

$$T_{1\to2} = -\int_{\mathbb{R}} (\mathcal{P}\rho)_2(x_2)\cdot\log(\mathcal{P}\rho)_2(x_2)\mathrm{d}x_2 = -\int_{\mathbb{R}} \frac{1}{b}\rho_1\left(\frac{x_2}{b}\right)\cdot\log\left[\frac{1}{b}\rho_1\left(\frac{x_2}{b}\right)\right]\mathrm{d}x_2 = \log b + H_1. \tag{46}$$

This is to say, the entropy transfer from the quadratic component (X_1) to the linear component (X_2) is the entropy of X_1 plus a part due to expansion/contraction of the total phase space. Particularly, when $b = 1$, i.e., when the volume of the phase space stays invariant, the transfer from X_1 to X_2 is just all that X_1 possesses.

This simple result of (46) is just what one would expect of the mapping component $\Phi_2(x_1, x_2) = bx_1$ in (42). As Φ_2 has no dependence on x_2 itself, the evolution of the entropy of X_2 is all through the transfer from the other component X_1. In other words, $T_{1\to2}$ is just the marginal entropy increase of X_2, which can be easily obtained, without going to the transfer formula (17). One may check that this is indeed equal to $\log b + H_1$. Among all the existing measures of information transfer, only our formalism yields a result for the Hénon map which is consistent with the dynamics (see (46)).

6. Summary

We have presented a rigorous formalism of information transfer for dynamical systems in terms of discrete mapping. Properties have been explored, and connections to a classical formalism established. A transfer measure has been obtained, validated, and used in two applications.

Information transfer occurs as a system evolves in time. It is measured by the amount of entropy transferred from one component to another in the course of this evolution. The rigorously derived transfer measure is unidirectional, namely, the information from X_i to X_j generally implies nothing about the information transfer in the other direction. Particularly, if X_i evolves independently of X_j ($i \neq j$), then the transfer from X_j to X_i is nil; while in the same time there could be a transfer in the other direction, should X_j relies X_i to grow. This is consistent with the property of transfer asymmetry emphasized by Schreiber [22].

The present formalism generalizes the 2D formalism in LK05 to maps with noninvertibility and systems with arbitrarily many dimensions. It is consistent with the 2D transfer measure previously obtained by different methods. It is also qualitatively consistent with the previously introduced transfer entropy of Schreiber.

We have applied the formalism to investigate the information transfers within the baker transformation and Hénon map. In both cases, results agree well with the facts observed or intuitively argued in the literature. For the Hénon map, our formalism yields a simple and physically clear transfer from the quadratic component to the linear component; for the baker transformation, we found that the stretching coordinate always loses information to the folding coordinate, while no transfer occurs in the opposite direction.

We close the paper by remarking that through taking limits we may use a discrete map to approach a continuous flow. In doing so a formalism of information transfer can be derived for continuous systems. This will be the subject of the second part of this study [15].

Acknowledgments

XSL has benefited from Andrew Majda's lectures on information theory. This work was supported by NSF under CMG Grant 0417728 to the Courant Institute of Mathematical Sciences.

Appendix. The Frobenius–Perron operator for the baker transformation and the dyadic transformation

For convenience, here we briefly summarize the derivation of the F–P operators for the baker transformation and the dyadic transformation. The material is based on [7,14].

A.1. The baker transformation

The baker transformation defined by (28) is invertible, and the inverse map is given by

$$\Phi^{-1}(x_1, x_2) = \begin{cases} \left(\frac{x_1}{2},\ 2x_2\right) & 0 \le x_2 \le \frac{1}{2},\ 0 \le x_1 \le 1 \\ \left(\frac{x_1+1}{2},\ 2x_2 - 1\right) & \frac{1}{2} \le x_2 \le 1,\ 0 \le x_1 \le 1. \end{cases} \quad \text{(A.1)}$$

Using Φ^{-1}, we can find the counterimage of $[0, x_1] \times [0, x_2]$ to be

(1) $0 \le x_2 < \frac{1}{2}$

$$\Phi^{-1}([0, x_1] \times [0, x_2]) = \left[0, \frac{x_1}{2}\right] \times [0, 2x_2];$$

(2) $\frac{1}{2} \le y \le 1$

$$\begin{aligned} \Phi^{-1}([0, x_1] \times [0, x_2]) &= \Phi^{-1}\left([0, x_1] \times \left[0, \frac{1}{2}\right]\right) \\ &\quad \cup \Phi^{-1}\left([0, x_1]\left[\frac{1}{2}, x_2\right]\right) \\ &= \left[0, \frac{x_1}{2}\right] \times [0, 1] \cup \left[\frac{1}{2}, \frac{x_1+1}{2}\right] \\ &\quad \times [0, 2x_2 - 1]. \end{aligned}$$

The F–P operator $\mathcal{P}$ is thus (see [14]):

$$\mathcal{P}\rho(x_1, x_2) = \frac{\partial^2}{\partial x_2 \partial x_1} \iint_{\Phi^{-1}([0,x_1]\times[0,x_2])} \rho(s, t)\, \mathrm{d}s\mathrm{d}t$$

which, after a series of transformations, gives

$$\mathcal{P}(x_1, x_2) = \begin{cases} \rho\left(\frac{x_1}{2}, 2x_2\right), & 0 \le x_2 < \frac{1}{2}, \\ \rho\left(\frac{1+x_1}{2}, 2x_2 - 1\right), & \frac{1}{2} \le x_2 \le 1. \end{cases} \quad \text{(A.2)}$$

A.2. The dyadic transformation

If only the stretching direction is considered, the baker transformation (28) is reduced to a dyadic transformation $\Phi : [0, 1] \mapsto [0, 1]$, $x_1 \mapsto 2x_1 \pmod 1$. It is easy to obtain

$$\Phi^{-1}([0, x_1]) = \left[0, \frac{x_1}{2}\right] \cup \left[\frac{1}{2}, \frac{1+x_1}{2}\right]$$

for $x_1 < 1$. So

$$\begin{aligned} \mathcal{P}\rho(x_1) &= \frac{\partial}{\partial x_1} \int_{\Phi^{-1}([0,x_1])} \rho(s)\mathrm{d}s \\ &= \frac{\partial}{\partial x_1} \int_0^{x_1/2} \rho(s)\mathrm{d}s + \frac{\partial}{\partial x_1} \int_{1/2}^{(1+x_1)/2} \rho(s)\,\mathrm{d}s \\ &= \frac{1}{2}\left[\rho\left(\frac{x_1}{2}\right) + \rho\left(\frac{1+x_1}{2}\right)\right]. \end{aligned}$$

References

[1] H.D.I. Abarbanel, Analysis of Observed Chaotic Data, Springer, New York, 1996.

[2] H.D.I. Abarbanel, N. Masuda, M.I. Rabinovich, E. Tumer, Distribution of mutual information, Phys. Lett. A 281 (2001) 368–373.

[3] R. Abramov, A.J. Majda, R. Kleeman, Information theory and predictability for low frequency variability, J. Atmos. Sci. 62 (2005) 65–87.

[4] J. Arnhold, et al., A robust method for detecting interdependences: Application to intracranially recorded EEG, Physica D 134 (1999) 419–430.

[5] T.M. Cove, J.A. Thomas, Elements of Information Theory, Wiley, NY, 1991.

[6] J.R. Dorfman, An Introduction to Chaos in Nonequilibrium Statistical Mechanics, Cambridge University Press, 1999.

[7] Pierre Gaspard, Chaos, Scattering and Statistical Mechanics, Cambridge University Press, 1998.

[8] A. Kaiser, T. Schreiber, Information transfer in continuous processes, Physica D 166 (2002) 43–62.

[9] K. Kaneko, Lyapunov analysis and information flow in coupled map lattices, Physica D 23 (1986) 436–447.

[10] Holger Kantz, Thomas Schreiber, Nonlinear Time Series Analysis, Cambridge University Press, 2004.

[11] R. Kleeman, Measuring dynamical prediction utility using relative entropy, J. Atmos. Sci. 59 (2002) 2057–2072.

[12] R. Kleeman, A.J. Majda, Predictability in a model of geostrophic turbulence, J. Atmos. Sci. 62 (2005) 2864–2879.

[13] R. Kleeman, Information flow in ensemble weather predictions, J. Atmos. Sci. 64 (2007) 1005–1016.

[14] A. Lasota, M.C. Mackey, Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics, Springer, New York, 1994.

[15] X.S. Liang, R. Kleeman, Information transfer between dynamical system components, Phys. Rev. Lett. 95 (24) (2005) 244101.

[16] X.S. Liang, R. Kleeman, A rigorous formalism of information transfer between dynamical system components. II. Continuous flow, Physica D 227 (2007) 173–182.

[17] A.J. Lichtenberger, M.A. Lieberman, Regular and Chaotic Dynamics, Springer, New York, 1992.

[18] R. Metzler, Y. Bar-Yam, M. Kardar, Information flow through a chaotic channel: Prediction and postdiction at finite resolution, Phys. Rev. E 70 (2004) 026205.

[19] Ernesto Pereda, Rodrigo Quian Quiroga, Joydeep Bhattacharya, Nonlinear multivariate analysis of neurophysiological signals, Prog. Neurobiol. (in press).

[20] M.G. Rosenblum, A.S. Pikovsky, J. Kurths, Phase synchronization of chaotic oscillators, Phys. Rev. Lett. 76 (1996) 1804–1807.

[21] T. Schreiber, Spatio-temporal structure in coupled map lattices: Two-point correlations versus mutual information, J. Phys. A 23 (1990) 393–398.

[22] T. Schreiber, Measuring information transfer, Phys. Rev. Lett. 85 (2) (2000) 461.

[23] J.A. Vastano, H.L. Swinney, Information transport in spatiotemporal systems, Phys. Rev. Lett. 60 (1988) 1773–1776.