Physica D 248 (2013) 1–15

Contents lists available at SciVerse ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Local predictability and information flow in complex dynamical systems

X. San Liang *

School of Marine Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China
School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China
China Institute for Advanced Study, Central University of Finance and Economics, Beijing 100081, China
Courant Institute of Mathematical Sciences, New York, NY 10012, USA

ARTICLE INFO

Article history:
Received 21 July 2012
Received in revised form
14 December 2012
Accepted 16 December 2012
Available online 5 January 2013
Communicated by T. Sauer

Keywords:
Information flow/transfer
Predictability
Causal relation
Turbulence closure
Hamiltonian
North Atlantic Oscillation

ABSTRACT

Predictability is by observation a local notion in complex dynamical systems. Its spatiotemporal structure implies a flow, or transfer in discrete cases, of information that redistributes the local predictability within the state space of concern. Information flow is a fundamental concept in general physics which has applications in a wide variety of disciplines such as neuroscience, material science, atmosphere–ocean science, and turbulence research, to name but a few. In this study, it is rigorously formulated with respect to relative entropy within the framework of a system with many components, each signifying a location or a structure. Given a component, the mechanism governing the evolution of its predictability can be classified into two groups, one due to the component itself, another due to a transfer of information from its peers. A measure of the transfer is rigorously derived, and an explicit expression obtained. This measure possesses a form reminiscent of that we have obtained before with respect to absolute entropy in [X.S. Liang and R. Kleeman, A rigorous formalism of information transfer between dynamical system components, Physica D 227 (2007) 173–182]; in particular, when the system is of dimensionality 2, there is no difference between the formalisms with respect to absolute entropy and relative entropy, except for a minus sign. Properties have been explored and discussed; particularly discussed is the property of asymmetry or causality, which states that information transfer from one component to another carries no hint about the transfer in the other direction, in contrast to the transfer of other quantities such as energy. This formalism has been applied to the study of the scale–scale interaction and information transfer between the first two modes of the truncated Burgers equation. It is found that all 12 transfers are essentially zero or negligible, save for a strong transfer between the sine components from the low-frequency mode to the high-frequency mode. That is to say, the predictability of the high-frequency mode is controlled by the knowledge of the low-frequency mode. This result, though from a highly idealized system, has interesting implications about the dynamical closure problem in turbulence research and atmosphere–ocean science, i.e., the subgrid processes may to some extent be parameterized by the large-scale dynamics. This study can be adopted to investigate the propagation of uncertainties in fluid flows, which has important applications in problems such as atmospheric observing platform design, and may be utilized to identify the route of information flowing within a complex network.

1. Introduction

In his pioneering work, Lorenz [1] shows that prediction of the state of a nonlinear dynamical system is impossible beyond a certain time limit if the system is intrinsically chaotic. This raises a severe issue in philosophy (e.g. [2]), and since then the problem of predictability has received enormous attention, in both theoretical dynamical systems (e.g., [3,4] and references therein) and applied fields such as atmosphere–ocean science (e.g., [5–20].) The past decades have seen a surge of interest in ensemble forecast (e.g., [21–31]); the fundamental scientific thrust is predictability.

Classically predictability is a global concept over the whole system. But in realistic problems, particularly in problems with high dimensional systems, people have observed that it generally varies from place to place. For example, Palmer [32] finds that his numerical weather model has different predictability for different flow regimes; Farrell [8] shows that predictability is structure dependent, and in the linear limit the most unpredictable structure can be identified; Kleeman [10] realizes the predictability difference between the El Niño modes; Tribbia [33] and Kleeman [10,12,25] have studied the predictability evolution in physical space. In other words, predictability is by observation a local concept, varying in physical space and/or phase space as it evolves in time.

* Correspondence to: School of Marine Sciences, Nanjing University of Information Science and Technology, 219 Ningliu Blvd, Nanjing 210044, China.
E-mail address: sanliang@courant.nyu.edu.

0167-2789/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2012.12.011

The spatiotemporal structure of predictability implies a flow, or transfer in discrete cases, of information that redistributes predictability from one place to another within the dynamical system of concern. This flow or transfer is important in that it determines how predictability in one place is altered due to other places, how uncertainties propagate in the system, and hence helps to identify the source region(s) of unpredictability. An immediate application is in observing platform design. In atmospheric science, for example, it has been argued that observations should target these source locations [25,33], in order for a weather forecast system to increase its forecast skill.

The above problem may be formalized within a framework of dynamical systems with many components, each component standing for a physical location or a structure. This way what we are discussing is essentially about the information transfer between dynamical system components, a concept which has been of interest for decades in communication, neuroscience, and nonlinear time series coherence analysis, to name but a few [34–42]. The available formalisms include the delayed mutual information [41] and the more sophisticated transfer entropy by Schreiber [40]. In relation to this study, these empirical/half-empirical formalisms have been applied to the investigation of the information flow in weather forecasts (e.g., [25]). Recently the notion of information transfer has been put on a rigorous footing by Liang and Kleeman in the context of dynamical systems (see [43–45], which hereafter will be cited as LK05, LK07a, LK07b, respectively). The resulting measure of the transfer is qualitatively consistent with the classical formalisms but is based on rigorous derivations. Explicit expressions have been obtained for continuous dynamical systems, and for two well-studied mappings, namely the baker transformation and the Hénon map. These results, most of them unique to the Liang–Kleeman formalism, agree well with what one may expect by physical intuition.

The Liang–Kleeman formalism is with respect to Shannon entropy, or absolute entropy as may appear in the literature. The predictability of a dynamical system, however, is measured by relative entropy. Relative entropy is also called Kullback–Leibler divergence; it is a measure of the difference between two probability distributions. Kleeman [10] points out that, in order to measure the utility of a forecast, one should ask how much additional information is added rather than how much information it has. Relative entropy provides a very natural measure of this information addition, if the reference probability is set as the initial distribution. Kleeman [10] also argues in favor of relative entropy because of its appealing properties, such as its invariance on nonlinear transformations and its non-negativity [46]. In the context of a Markov chain, it has been proved that it always decreases monotonically with time, a property usually referred to as the generalized second law of thermodynamics (e.g., [46]; also see [47]). The concept of relative entropy is now a well accepted measure of predictability (e.g., [48,10,49]).

Our problem here is therefore fundamentally how information is transferred with respect to relative entropy. The purpose of this study is to develop a formalism for this information transfer. The development parallels what we have done in LK07b, and the resulting transfer will be referred to as *information transfer with respect to relative entropy*, or simply *information transfer* when no confusion arises. In the literature the term "information flow" is also seen (e.g., [25,50,33]) for the same meaning. We recognize that it might be more appropriate to use "transfer" for discrete systems, but it indeed forms a flow when the system components are associated with locations in physical space. We will hence use the two in the text without distinction.

In the following we first present a conceptual framework for this study, then give the concept a formal definition. An explicit formula is derived for the information transfer of concern; the detailed derivations are supplied in Sections 3–5. Properties of the formulation are investigated and discussed (Section 6), and an application presented (Section 7). This study is summarized in Section 8.

2. Mathematical framework

Information flow or information transfer is a fundamental concept in general physics and dynamical system which has a wide variety of applications in natural and social sciences (e.g., [37,25,51,52,33]). However, for decades there have been available in the community only empirical or half-empirical formalisms, e.g., the time-delayed information transfer [41] and the transfer entropy [40] with respect to a Markov chain. These formalisms, albeit different in form, all deal with the time evolution of entropy, as observed in LK07a. This observation motivates Liang and Kleeman to build a unified formalism, which later on they find can be put on a rigorous footing within the framework of a dynamical system (LK05, LK07a,b). In the following, we will follow the same route to build the formalism for this study.

Consider an n-dimensional dynamical system:

$$\frac{dx_1}{dt} = F_1(t; x_1, x_2, \ldots, x_n), \tag{1}$$

$$\frac{dx_2}{dt} = F_2(t; x_1, x_2, \ldots, x_n), \tag{2}$$

$$\vdots \qquad \vdots$$

$$\frac{dx_n}{dt} = F_n(t; x_1, x_2, \ldots, x_n). \tag{3}$$

for state variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. We want to understand how the predictability of one component of $\mathbf{x}$ is altered by another, namely, how information is transferred between two components with respect to relative entropy. For simplicity, the above equation set may also appear in the text in a vectorial form,

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(t; \mathbf{x}), \tag{4}$$

with the flow $\mathbf{F} = (F_1, F_2, \ldots, F_n)$. Denote by $\mathbf{X} = (X_1, X_2, \ldots, X_n) \in \Omega$ the random variables corresponding to $(x_1, x_2, \ldots, x_n)$, where Ω is the sample space, and let $\rho = \rho(t; x_1, x_2, \ldots, x_n)$ be the joint probability density of $\mathbf{X}$. Assume

$$\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n, \tag{5}$$

(Ω_i is the sample spaces of X_i, $i = 1, 2, \ldots, n$), and write

$$\Omega_{jn} \equiv \Omega_j \times \Omega_{j+1} \times \cdots \times \Omega_n, \quad j = 1, 2, \ldots, n-1 \tag{6}$$

throughout for notational convenience. Further assume that ρ vanishes at the boundaries of Ω, i.e., the extreme events have a measure of zero in the probability space. These assumptions have been justified for real problems in our previous studies (LK05, LK07b). For many problems of interest, $\Omega = \mathbb{R}^n$. In this case, the assumption of vanishing ρ at boundaries automatically holds, since ρ is compactly supported.

In LK05, LK07a,b, a rigorous formalism of information transfer between the components is developed with respect to absolute entropy

$$H = -\int_\Omega \rho \log \rho \, d\mathbf{x}. \tag{7}$$

From the Liouville equation (e.g., [53]) associated with (1)–(3) or (4), a remarkable law was obtained in LK05 that governs the time evolution of H,

$$\frac{dH}{dt} = E(\nabla \cdot \mathbf{F}) = \sum_{i=1}^{n} E\left(\frac{\partial F_i}{\partial x_i}\right), \tag{8}$$

where E stands for the operator of mathematical expectation with respect to density ρ. (Refer to Section 3 where a proof will be given.) Eq. (8) reads that, given a dynamical system, the rate of change of absolute entropy is simply the expectation of the divergence of the flow with the system, and contributions from different components add up to make the change. In LK05, this fact is used to establish the information transfer formalism. Without loss of generality, consider only the transfer from X_2 to X_1; if not, the variables may always be reordered to make it so. We thence need the marginal entropy of X_1:

$$H_1 = -\int_{\Omega_1} \rho_1 \log \rho_1 \, dx_1, \tag{9}$$

where

$$\rho_1 = \int_{\Omega_{2n}} \rho \, dx_2 dx_3 \dots dx_n, \tag{10}$$

is the marginal density. The time rate of change of H_1, namely $\frac{dH_1}{dt}$, has contributions from two different dynamical sources: one with the effect from X_2 excluded, another one from X_2. The latter is precisely what we are looking for, i.e., the information transfer from X_2 to X_1 with respect to absolute entropy. The former part needs some clarification. For a system with components $(X_1, X_2, X_3, \dots, X_n)$, that "the effect of X_2 is excluded" means that the corresponding deterministic variable x_2 stays fixed, entering the system as a parameter. In the present context, this means a modification of the system (1)–(3) to

$$\frac{dx_1}{dt} = F_1(t; x_1, x_2, x_3, \dots, x_n),$$
$$\frac{dx_3}{dt} = F_3(t; x_1, x_2, x_3, \dots, x_n),$$
$$\vdots \qquad \vdots$$
$$\frac{dx_n}{dt} = F_n(t; x_1, x_2, x_3, \dots, x_n),$$

where the second equation is excluded, and x_2 now appears as a parameter. Write this part as $dH_{1\not{2}}/dt$, with subscript $\not{2}$ signifying that the effect of X_2 is excluded, then the transfer is $dH_1/dt - dH_{1\not{2}}/dt$. Here dH_1/dt can be easily derived from the Liouville equation; the key is to find $dH_{1\not{2}}/dt$. For a 2D system, i.e., when $n = 2$, $dH_{1\not{2}}/dt$ is identical to the time rate of change of the absolute entropy of X_1 as X_1 evolves on its own. Based on the observation about (8), we then have, for $n = 2$,

$$\frac{dH_{1\not{2}}}{dt} = E\left(\frac{\partial F_1}{\partial x_1}\right). \tag{11}$$

Eq. (11) was originally obtained in LK05 through an intuitive argument, and was later on proved in LK07b. Accordingly the information transfer with respect to absolute entropy is obtained by subtracting $dH_{1\not{2}}/dt$ from dH_1/dt. This formalism is rigorous in nature, and has been validated with benchmark dynamical system problems like the baker transformation and Hénon map.

We here need to develop another formalism of information flow, following the same route of development as above, but with respect to relative entropy

$$D = D(\rho \parallel q) = \int_\Omega \rho \log \frac{\rho}{q} d\mathbf{x} = -H - \int_\Omega \rho \log q \, d\mathbf{x}. \tag{12}$$

In the definition, q is a density at some fixed time. Usually it is the initial density or density in the equilibrium state; here let it be the initial density to avoid any confusion that may arise. With respect to D, we are concerned with the information transfer between two components. Again without loss of generality, consider only the transfer from X_2 to X_1. We hence need the marginal relative entropy of X_1:

$$D_1 = \int_{\Omega_1} \rho_1 \log \frac{\rho_1}{q_1} dx_1 = -H_1 - \int_{\Omega_1} \rho_1 \log q_1 \, dx_1, \tag{13}$$

where

$$q_1 = \int_{\Omega_{2n}} q \, dx_2 dx_3 \dots dx_n. \tag{14}$$

Following the above argument, the mechanisms governing the time evolution of D_1 can be classified exclusively into two groups: one from a modified system with the effect of X_2 excluded, another from the component X_2. This latter mechanism is what we are seeking for, namely, the information transfer from X_2 to X_1. Correspondingly $\frac{dD_1}{dt}$, the time rate of change of D_1, can be decomposed as the change of D_1 with the effect of X_2 instantaneously excluded, denoted as $\frac{dD_{1\not{2}}}{dt}$, plus the time rate of information transfer, $T_{2\to1}$. So

$$T_{2\to1} = \frac{dD_1}{dt} - \frac{dD_{1\not{2}}}{dt}. \tag{15}$$

The units of $T_{2\to1}$ vary, depending on the base of the logarithm in D_1 that is used. The common units are nats per second for base e and bits per second for base 2.

The above formalism allows a clear interpretation for information flow/transfer. In the present context, the transfer from X_2 to X_1, i.e., $T_{2\to1}$, gives quantitatively the effect of X_2 on the predictability of X_1. Particularly, a positive $T_{2\to1}$ means that the evolution of X_2 favors the prediction of X_1; in other words, it will make X_1 more predictable. On the other hand, a negative $T_{2\to1}$ implies that X_2 reduces the predictability of X_1. (Previously in LK07a and LK07b when Shannon entropy is considered, a negative transfer $T_{2\to1}$ means that the evolution of X_2 tends to reduce the uncertainty of X_1, and vice versa.)

The whole problem is now converted to the derivation of $\frac{dD_1}{dt}$ and $\frac{dD_{1\not{2}}}{dt}$. In the following section $\frac{dD_1}{dt}$ is derived. As in LK07b, the challenge comes from the evaluation of $D_{1\not{2}}$, which we defer to Section 4. For convenience, the notation $\not{2}$ in the subscript of $D_{1\not{2}}$ will be extended to any $\not{j}$ to signify that component j (or the effect of component j) is excluded from a set of n independent variables. For example,

$$\rho_{\not{2}} = \rho_{\not{2}}(x_1, x_3, \dots, x_n) = \int_{\Omega_2} \rho(\mathbf{x}) \, dx_2, \tag{16}$$

$$\rho_{\not{1}\not{2}} = \rho_{\not{1}\not{2}}(x_3, \dots, x_n) = \int_{\Omega_1 \times \Omega_2} \rho(\mathbf{x}) \, dx_1 dx_2, \tag{17}$$

where the time dependency has been suppressed for notation simplicity. This convention will be used throughout the paper without further clarification.

3. Time rate of change of D_1

Theorem 1. *For the system described in Section 2,*

$$\frac{dH}{dt} = E(\nabla \cdot \mathbf{F}), \tag{18}$$

$$\frac{dD}{dt} = -E(\nabla \cdot \mathbf{F}) - E(\mathbf{F} \cdot \nabla \log q) \tag{19}$$

$$\frac{dH_1}{dt} = \int_\Omega \log \rho_1 \cdot \frac{\partial (F_1 \rho)}{\partial x_1} \, d\mathbf{x}, \tag{20}$$

$$\frac{dD_1}{dt} = -\int_\Omega \log \frac{\rho_1}{q_1} \cdot \frac{\partial (F_1 \rho)}{\partial x_1} \, d\mathbf{x}, \tag{21}$$

where E is the operator of expectation with respect to the density ρ.

Proof. See Appendix A. □

Eq. (18) was obtained in LK05; it states that the change of absolute entropy in a system is totally controlled by the divergence of the flow, or the contraction/expansion of the phase space. Notice that $\int_\Omega \frac{\partial(F_1\rho)}{\partial x_1}\, d\mathbf{x} = 0$ by the assumptions introduced before (vanishing density at boundaries and Cartesian product form for Ω). So Eq. (21) can be equivalently written as

$$\frac{dD_1}{dt} = -\int_\Omega \left(1 + \log\frac{\rho_1}{q_1}\right) \cdot \frac{\partial(F_1\rho)}{\partial x_1}\, d\mathbf{x}. \tag{22}$$

Later on we will have opportunity to use (22).

4. Time rate of change of D_1 with the effect of X_2 excluded

Theorem 2. *For the dynamical system described in Section 2, the time rate of change of the marginal relative entropy of X_1 with the effect of X_2 excluded is*

$$\frac{dD_{1\not 2}}{dt} = -\int_\Omega \left(1 + \log\frac{\rho_1}{q_1}\right) \frac{\partial(F_1\rho_{\not 2})}{\partial x_1}\Theta_{2|1}\, d\mathbf{x} + \int_\Omega \frac{\partial\left(F_1\rho_1\log\frac{\rho_1}{q_1}\right)}{\partial x_1}\theta_{2|1}\, d\mathbf{x}, \tag{23}$$

where

$$\theta_{2|1} = \theta_{2|1}(x_1, x_2, x_3, \ldots, x_n) = \frac{\rho}{\rho_{\not 2}}\rho_{\not 1\not 2}, \tag{24}$$

$$\Theta_{2|1} = \Theta_{2|1}(x_1, x_2) = \int_{\Omega_{3n}} \theta_{2|1}\, dx_3 \ldots dx_n. \tag{25}$$

This theorem cannot be proved using the Liouville equation corresponding to (4), as the dynamics is changed upon manipulating x_2. In LK07b, the problem is approached by discretizing (1)–(3) or (4), finding how H_1 increases from time t to time $t + \Delta t$ in the absence of the influence of X_2, and then taking the limit as $\Delta t \to 0$. In the following the same strategy is adopted.

Discretization of the continuous system (4) results in a mapping $\Phi : \Omega \longrightarrow \Omega, \mathbf{x} \mapsto \mathbf{y}$ such that

$$\Phi : \mathbf{y} = \mathbf{x} + \Delta t\, \mathbf{F}(t; \mathbf{x}), \tag{26}$$

i.e., an approximation of (4) up to the first order of Δt. To avoid confusion, here $\mathbf{x}(t+\Delta t)$ has been written as $\mathbf{y} = (y_1, y_2, \ldots, y_n)$; this convention will be kept throughout. In component form, the mapping is

$$\Phi = (\Phi_1, \Phi_2, \ldots, \Phi_n) : \begin{cases} y_1 = x_1 + \Delta t \cdot F_1(t; \mathbf{x}), \\ y_2 = x_2 + \Delta t \cdot F_2(t; \mathbf{x}), \\ \vdots \quad \vdots \\ y_n = x_n + \Delta t \cdot F_n(t; \mathbf{x}). \end{cases} \tag{27}$$

Corresponding to Φ that maps the state from t to $t + \Delta t$, there is an operator sending the density of the state variables from t to $t + \Delta t$. This is the Frobenius–Perron operator, or F–P operator for short. Formally, the F–P operator, written as $\mathscr{P}$, corresponding to a transformation $\Phi : \Omega \mapsto \Omega$ is a map $\mathscr{P} : L^1(\Omega) \mapsto L^1(\Omega)$ such that, for any $\omega \subset \Omega$,

$$\int_\omega \mathscr{P}\rho(\mathbf{x})\, d\mathbf{x} = \int_{\Phi^{-1}(\omega)} \rho(\mathbf{x})\, d\mathbf{x}.$$

It can be viewed as the discrete form of the Liouville equation for density ρ. See [53] for more details. Liang and Kleeman (LK07b) have shown that the mapping Φ and its associated F–P operator $\mathscr{P}$ possess some interesting properties, which here we briefly summarize.

(1) As Δt goes to zero, Φ and its individual components are always invertible, and

$$\Phi^{-1} : \begin{cases} x_1 = y_1 - \Delta t \cdot F_1(t; \mathbf{y}) + O(\Delta t^2), \\ x_2 = y_2 - \Delta t \cdot F_2(t; \mathbf{y}) + O(\Delta t^2), \\ \vdots \quad \vdots \\ x_n = y_n - \Delta t \cdot F_n(t; \mathbf{y}) + O(\Delta t^2). \end{cases} \tag{28}$$

(2) The Jacobian of Φ, J, and its inverse, are

$$J = \det\left[\frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)}\right] = 1 + \Delta t \nabla\cdot\mathbf{F} + O(\Delta t^2); \tag{29}$$

$$J^{-1} = 1 - \Delta t \nabla\cdot\mathbf{F} + O(\Delta t^2). \tag{30}$$

(3) The F–P operator $\mathscr{P}$ can be explicitly written out:

$$\mathscr{P}\rho(y_1, \ldots, y_n) = \rho\left(\Phi^{-1}(y_1, \ldots, y_n)\right) \left|J^{-1}\right| = \rho(x_1, x_2, \ldots, x_n) \left|J^{-1}\right|, \tag{31}$$

thanks to the invertibility of Φ (cf. [53]).

Following LK07b, in order to get rid of the effect of X_2, we need to freeze x_2 at t. The mapping Φ is accordingly modified, with the second equation removed from the set, and a new transformation results:

$$\Phi_{\not 2} : \begin{cases} y_1 = x_1 + \Delta t \cdot F_1(t; \mathbf{x}), \\ y_3 = x_3 + \Delta t \cdot F_3(t; \mathbf{x}), \\ \vdots \quad \vdots \\ y_n = x_n + \Delta t \cdot F_n(t; \mathbf{x}). \end{cases} \tag{32}$$

$\Phi_{\not 2}$ maps $(x_1, x_3, x_4, \ldots, x_n)$ to $(y_1, y_3, y_4, \ldots, y_n)$ with x_2 as a parameter. Corresponding to $\Phi_{\not 2}$ there is an F–P operator. Write it as $\mathscr{P}_{\hat 2}$. Here we use the subscript $\hat 2$ to signify that component x_2 is frozen; we do not use $\not 2$ because in the resulting probability density, x_2 is not excluded, but retained as a parameter. (Think about the probability conditioned on X_2.) We thence have $\mathscr{P}_{\hat 2}\rho$, the joint density at time $t + \Delta t$ with x_2 frozen as a parameter at time t, and

$$(\mathscr{P}_{\hat 2}\rho)_1(y_1) = \int_{\Omega_{3n}} \mathscr{P}_{\hat 2}\rho(y_1, y_3, \ldots, y_n)\, dy_3 \ldots dy_n$$

being the corresponding marginal density of $Y_1 = X_1(t + \Delta t)$. (Again note $(\mathscr{P}_{\hat 2}\rho)_1$ has dependence on x_2.) From LK07a and LK07b, the marginal absolute entropy for X_1 evolved from H_1 with contribution from X_2 excluded since time t is

$$H_{1\not 2}(t + \Delta t) = \int_\Omega (\mathscr{P}_{\hat 2}\rho)_1(y_1) \log(\mathscr{P}_{\hat 2}\rho)_1(y_1) \cdot\rho(x_2|x_1, x_3, \ldots, x_n) \cdot\rho_{3\ldots n}(x_3, \ldots, x_n)\, dy_1 dx_2 \ldots dx_n,$$

where $y_1 = x_1 + \Delta t\, F_1(\mathbf{x})$, $\rho_{3\ldots n} = \rho_{\not 1\not 2}$, $\rho(x_2|x_1, x_3, \ldots, x_n)$ being the conditional density of X_2 on $(X_1, X_3, \ldots, X_n)$. By the same argument, we have

$$D_{1\not 2}(t + \Delta t) = \int_\Omega (\mathscr{P}_{\hat 2}\rho)_1(y_1) \log\frac{(\mathscr{P}_{\hat 2}\rho)_1(y_1)}{q_1(y_1)} \cdot\rho(x_2|x_1, x_3, \ldots, x_n) \cdot\rho_{3\ldots n}(x_3, \ldots, x_n)\, dy_1 dx_2 \ldots dx_n. \tag{33}$$

To evaluate $D_{1\not 2}(t + \Delta t)$, the key is the evaluation of the F–P operator associated with the modified mapping $\Phi_{\not 2}$:

Proposition 3.

$$(\mathscr{P}_{\hat 2}\rho)_1(y_1) = \rho_1(y_1) - \Delta t \cdot \int_{\Omega_{3n}} \frac{\partial F_1\rho_{\not 2}}{\partial y_1}\, dx_3 \ldots dx_n + O(\Delta t^2). \tag{34}$$

Proof. See Appendix B. With this proposition, we are ready to prove the Theorem of this section. □

Outline of the proof of Theorem 2

Substitute (34) into (33), and express the x_1 in $\rho(x_2|x_1, x_3, \ldots, x_n)$ as a function of y_1 (from the inverse map $\Phi_{\not 2}^{-1}$). Then compute $\frac{D_{1\not 2}(t+\Delta t)-D_1(t)}{\Delta t}$. Eq. (23) follows as $\Delta t \to 0$. Refer to Appendix C for the lengthy detailed derivation.

5. Information transfer with respect to relative entropy

Theorem 4. *For the system described in Section 2, the information transfer with respect to relative entropy from X_2 to X_1 is*

$$T_{2\to 1} = -\int_\Omega \left(1+\log\frac{\rho_1}{q_1}\right)\cdot\left(\frac{\partial F_1\rho}{\partial x_1} - \frac{\partial F_1 \rho_{\not 2}}{\partial x_1}\Theta_{2|1}\right) d\mathbf{x} - \int_\Omega \frac{\partial}{\partial x_1}\left(F_1\rho_1 \log\frac{\rho_1}{q_1}\right)\cdot \theta_{2|1}\, d\mathbf{x}, \tag{35}$$

where

$$\theta_{2|1} = \theta_{2|1}(x_1, x_2, x_3, \ldots, x_n) = \frac{\rho}{\rho_{\not 2}}\rho_{\not 1\not 2},$$

$$\Theta_{2|1} = \Theta_{2|1}(x_1, x_2) = \int_{\Omega_{3n}} \theta_{2|1}\, dx_3 \ldots dx_n,$$

and $\Theta_{2|1}$ may be viewed as a generalized conditional density of X_2 on X_1.

Proof. Subtract (23) from (22) and the result follows. □

Note the transfer rate of relative entropy is in a form similar to that for absolute entropy. One changes $\log\rho_1$ in the Eq. (51) of LK07b into $\log\frac{\rho_1}{q_1}$ and multiplies the whole formula by (-1), and obtains Eq. (35). Recalling the definition of relative entropy (12), this is just what one may expect.

Above is the transfer from X_2 to X_1. Following the same procedure, it is easy to arrive at the transfer from X_j to X_i, for any $i, j = 1, 2, \ldots, n$, $i \neq j$. One may replace the index 2 by j, and 1 by i in (35), and make the corresponding modification for $\theta_{2|1}$ and $\Theta_{2|1}$ to obtain the formula. But the easiest way is to rearrange the order of (1)–(3) such that j is in the second slot and i in the first. This way the rate of transfer is expressed in the same form as (35).

6. Properties

The information transfer formulated in (35) possesses some nice properties. The first is about *causality* or *transfer asymmetry*. In physics, the transfer of energy (or other physical properties such as mass) between two entities is usually anti-symmetric. That is to say, if entity A transfers an amount of energy, say, ΔE, to entity B, then by the law of energy conservation it is equivalent to saying that B transfers $-\Delta E$ to A. (This simple fact has been used to formulate a new localized hydrodynamic instability analysis [54].) In particular, when no energy transfer occurs in a direction, the transfer in the opposite direction also vanishes. Information transfer or information flow, however, behaves quite differently. The transfers $T_{2\to 1}$ and $T_{1\to 2}$ from (35) between X_1 and X_2 are usually not related; in other words, the transfer is asymmetric or directional. This is a fundamental property for information transfer; its importance lies in its implication of causality, as elaborated in the literature [40]. The following theorem is a concretization of it.

Theorem 5 (*Causality*)**.** *For the system introduced in Section 2, if F_1 is independent of x_2, then $T_{2\to 1} = 0$. In the meantime, $T_{1\to 2}$ does not need to be zero, unless F_2 is independent of x_1.*

Proof. If F_1 is independent of x_2, in (35) the ρ, $\Theta_{2|1}$, and $\theta_{2|1}$ can be integrated separately with respect to x_2. Observe that

$$\int \rho\, dx_2 = \rho_{\not 2},$$

$$\int \theta_{2|1} dx_2 = \int \frac{\rho}{\rho_{\not 2}}\rho_{\not 1\not 2}\, dx_2 = \rho_{\not 1\not 2},$$

$$\int \Theta_{2|1} dx_2 = \int\left(\int \theta_{2|1} dx_3\ldots dx_n\right) dx_2 = \int \rho_{\not 1\not 2}\, dx_3 \ldots dx_n = 1.$$

So integrating (35) once with respect to x_2 gives

$$\begin{aligned} T_{2\to 1} &= -\int\left(1+\log\frac{\rho_1}{q_1}\right) \cdot\left(\frac{\partial F_1\rho_{\not 2}}{\partial x_1} - \frac{\partial F_1\rho_{\not 2}}{\partial x_1}\right) dx_1 dx_3\ldots dx_n \\ &\quad - \int \frac{\partial}{\partial x_1}\left(F_1\rho_1\log\frac{\rho_1}{q_1}\right)\cdot \rho_{\not 1\not 2}\, dx_1 dx_3 \ldots dx_n \\ &= -\int \frac{\partial}{\partial x_1}\left(F_1\rho_1 \log\frac{\rho_1}{q_1}\rho_{\not 1\not 2}\right) dx_1 dx_3\ldots dx_n \\ &= 0. \quad \square \end{aligned} \tag{36}$$

Another property relates Theorem 2 to (11), the 2D results of LK05:

Theorem 6. *When $n = 2$, $\frac{dD_{1\not 2}}{dt} = -E\left(\frac{\partial F_1}{\partial x_1}\right) - E\left(F_1\frac{\partial \log q_1}{\partial x_1}\right)$.*

Proof. By definition, when $n = 2$,

$$\Theta_{2|1} = \theta_{2|1} = \rho(x_2|x_1) = \frac{\rho}{\rho_1},$$

and

$$\rho_{\not 2} = \rho_1.$$

So

$$\begin{aligned} \frac{dD_{1\not 2}}{dt} &= -\int_\Omega\left(1+\log\frac{\rho_1}{q_1}\right)\frac{\partial F_1\rho_{\not 2}}{\partial x_1}\Theta_{2|1}\, d\mathbf{x} + \int_\Omega \frac{\partial\left(F_1\rho_1\log\frac{\rho_1}{q_1}\right)}{\partial x_1}\theta_{2|1}\, d\mathbf{x} \\ &= -\int_\Omega\left(1+\log\frac{\rho_1}{q_1}\right)\frac{\partial F_1\rho_1}{\partial x_1}\frac{\rho}{\rho_1}\, d\mathbf{x} + \int_\Omega\frac{\partial\left(F_1\rho_1\log\frac{\rho_1}{q_1}\right)}{\partial x_1}\frac{\rho}{\rho_1}\, d\mathbf{x} \\ &= \int_\Omega\left[-\rho\frac{F_1}{q_1}\frac{\partial q_1}{\partial x_1} - \rho\frac{\partial F_1}{\partial x_1}\right] d\mathbf{x} \\ &= -E\left(\frac{\partial F_1}{\partial x_1}\right) - E\left(F_1\frac{\partial\log q_1}{\partial x_1}\right). \quad \square \end{aligned}$$

In this theorem, the remarkable LK05 result (11) is recovered in the first term on the right hand side; the second term arises from the reference entropy q. The existence of the second term makes the present formalism generally different from that of LK05. But, surprisingly, the 2D information transfers obtained from these different formalisms are precisely the same, save for a minus sign which is due to the different definitions of D and H. Indeed, corresponding to Eq. (10) of LK05, we have

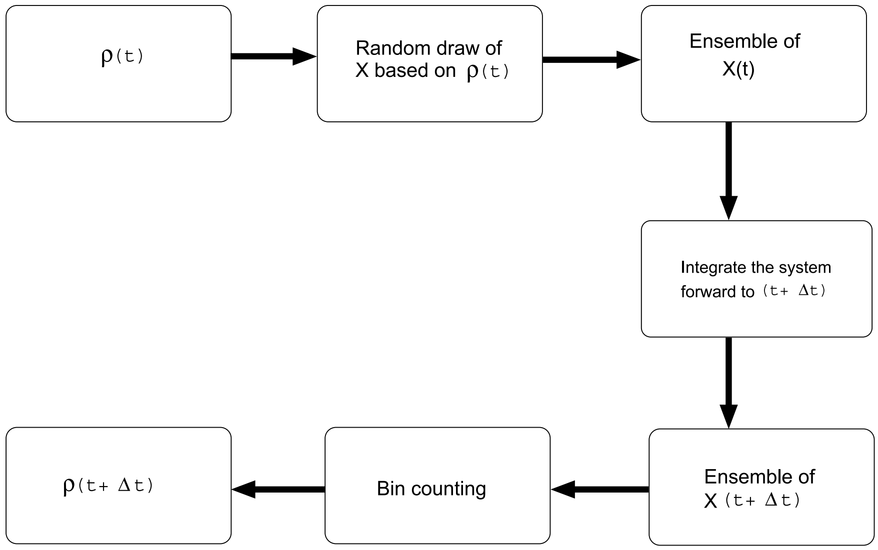


Fig. 1. Numerical scheme to solve the joint density ρ. Instead of obtaining $\rho(t+\Delta t)$ directly from $\rho(t)$ by solving the Liouville equation, a detour is made through ensemble prediction of the TBS system.

Corollary 7. *When* $n = 2$, (35) *becomes*

$$T_{2\to1} = E_{2|1}\left[\frac{\partial(F_1\rho_1)}{\partial x_1}\right], \tag{37}$$

where $E_{2|1}$ *stands for the expectation with respect to conditional density* $\rho_{2|1} = \rho/\rho_1$.

Proof. As shown above, when $n = 2$, $\rho_{\not 2} = \rho_1, \theta_{2|1} = \Theta_{2|1} = \rho/\rho_1 = \rho_{2|1}$. Substituting these back to (35), and using the fact that ρ is compactly supported, we have

$$\begin{aligned}
T_{2\to1} &= -\int_\Omega \left(1+\log\frac{\rho_1}{q_1}\right)\cdot\left(\frac{\partial F_1\rho}{\partial x_1} - \frac{\partial F_1\rho_1}{\partial x_1}\frac{\rho}{\rho_1}\right)d\mathbf{x} \\
&\quad - \int_\Omega \frac{\partial\left(F_1\rho_1\log\frac{\rho_1}{q_1}\right)}{\partial x_1}\cdot\frac{\rho}{\rho_1}\,d\mathbf{x} \\
&= -\int_\Omega \frac{\partial(F_1\rho_1)}{\partial x_1}\,d\mathbf{x} + \int_\Omega \frac{\partial(F_1\rho_1)}{\partial x_1}\cdot\frac{\rho}{\rho_1}\,d\mathbf{x} \\
&\quad + \int_\Omega \left[-\log\frac{\rho_1}{q_1}\cdot\frac{\partial F_1\rho}{\partial x_1} + \log\frac{\rho_1}{q_1}\cdot\frac{\partial F_1\rho_1}{\partial x_1}\frac{\rho}{\rho_1}\right. \\
&\quad \left. - \frac{\partial\left(F_1\rho_1\log\frac{\rho_1}{q_1}\right)}{\partial x_1}\frac{\rho}{\rho_1}\right]d\mathbf{x} \\
&= \int_\Omega \frac{\partial(F_1\rho_1)}{\partial x_1}\frac{\rho}{\rho_1}\,d\mathbf{x} - \int_\Omega \frac{\partial\left(F_1\rho\log\frac{\rho_1}{q_1}\right)}{\partial x_1}\,d\mathbf{x} \\
&= \int_\Omega \frac{\partial(F_1\rho_1)}{\partial x_1}\rho_{2|1}\,d\mathbf{x} = E_{2|1}\left[\frac{\partial(F_1\rho_1)}{\partial x_1}\right]. \quad \square
\end{aligned}$$

7. Application with the truncated Burgers–Hopf system

As an example of application, we re-consider the truncated Burgers–Hopf system (TBS) examined in LK07b, a system first introduced in [55,56] to explore a stochastic scheme of parameterization of the unresolved processes in numerical weather forecasts. It is obtained through a Galerkin truncation of the inviscid Burgers equation. If only two modes are retained, we have the following four-dimensional autonomous system (see LK07b):

$$\frac{dx_1}{dt} = F_1(x_1,x_2,x_3,x_4) = x_1x_4 - x_3x_2 \tag{38}$$

$$\frac{dx_2}{dt} = F_2(x_1,x_2,x_3,x_4) = -x_1x_3 - x_2x_4, \tag{39}$$

$$\frac{dx_3}{dt} = F_3(x_1,x_2) = 2x_1x_2, \tag{40}$$

$$\frac{dx_4}{dt} = F_4(x_1,x_2) = -x_1^2 + x_2^2, \tag{41}$$

where (x_1, x_2) are the cosine and sine components of the first mode, and (x_3, x_4) the components of the second mode, respectively. The TBS is intrinsically chaotic with a low-dimensional attractor [55–57]. For this particular case, there are two linearly independent first integrals, namely energy and Hamiltonian, and with them the system becomes integrable through a canonical transformation of coordinates; see Appendix D for details. We now study the information transfers with respect to relative entropy between the four components, and compare the results to those in LK07b.

The key to the computation of the information transfer (35) is the estimation of the joint density of (X_1, X_2, X_3, X_4) as a function of time. This may be obtained through solving

$$\frac{\partial\rho}{\partial t} + \frac{\partial(F_1\rho)}{\partial x_1} + \frac{\partial(F_2\rho)}{\partial x_2} + \frac{\partial(F_3\rho)}{\partial x_3} + \frac{\partial(F_4\rho)}{\partial x_4} = 0, \tag{42}$$

the Liouville equation corresponding to Eqs. (38)–(41). A more efficient way is, instead of solving ρ directly, estimating the density with the ensembles generated at each time step from ensemble prediction of (38)–(41), as schematized in Fig. 1. The ensemble is formed with the trajectories randomly distributed in the beginning, through solving (38)–(41) using the second order Runge–Kutta method with a time step size $\Delta t = 0.01$. A typical computed trajectory is plotted in Figs. 2 and 3; it shows an invariant manifold or strange attractor limited within some finite domain. If we write $\Omega_d = [-d, d] \times [-d, d] \times [-d, d] \times [-d, d]$, the strange attractor lies within Ω_{25}. So a finite domain may be chosen for the computation, although ρ is defined on $\mathbb{R}^4$. We choose it to be Ω_{30}, a domain slightly larger than Ω_{25}. This domain is then uniformly partitioned into $30 \times 30 \times 30 \times 30$ bins, with a spacing of 2 in each dimension. A huge ensemble of initial conditions of size

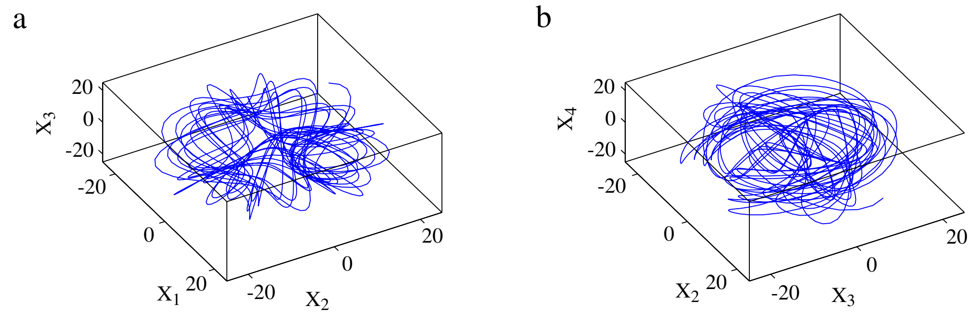


Fig. 2. A trajectory of (38)–(41) starting from $t = 0$ at $\mathbf{x}(0) = (40, 40, 40, 40)$. It is attracted into a domain as shown after $t = 1.5$ (the part prior to $t = 1.5$ not plotted). (a) and (b) are projections in subspaces x_1-x_2-x_3 and x_2-x_3-x_4, respectively.

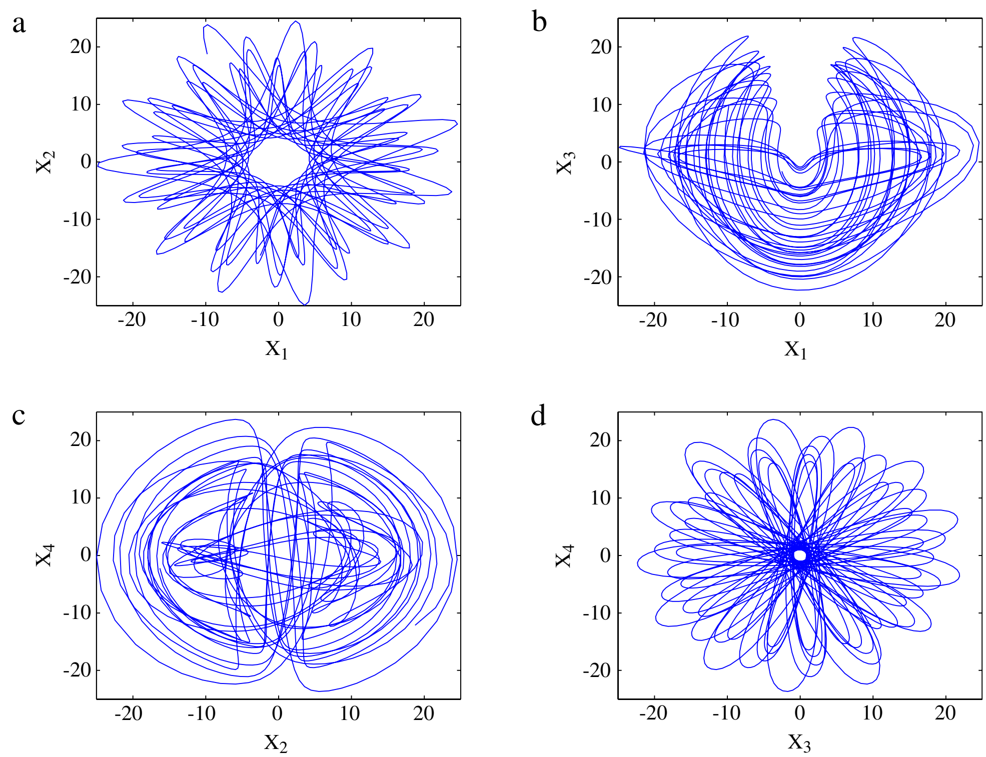


Fig. 3. As Fig. 2, but for projections on 2D planes.

$40^4 = 2.56 \times 10^6$ is first generated, through drawing randomly according to a preset distribution $\rho_0(\mathbf{x})$. We adopt an ensemble size of 40^4 instead of 30^4 to ensure more than one draw per bin on average. Suppose $\mathbf{x}$ is initially distributed as a Gaussian $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with a mean $\boldsymbol{\mu} = (\mu_i)$ and a covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$, $i, j = 1, 2, 3, 4$, and suppose $\sigma_{ii} = \sigma_i^2$, and $\sigma_{ij} = 0$ if $i \neq j$. The parameters μ_i and $\sigma_i^2 (i = 1, 2, 3, 4)$ are open for experiments. With these initial conditions, Eqs. (38)–(41) are integrated forward. At every time step we obtain an ensemble of $\mathbf{X}$, and therefrom a joint density ρ through bin-counting.

The transfers $T_{j\to i}$, $i, j = 1, 2, 3, 4$, $i \neq j$, now can be computed straightforwardly by evaluating (35). As in LK07b, there are 12 series to compute. Notice that in (40) and (41), the evolutions of x_3 and x_4 do not depend on x_3 and x_4, so $T_{3\to 4} = T_{4\to 3} = 0$ by Theorem 5. The computed results agree with this inference. The other transfers have only numerical solutions, and may vary with the initial distribution. We have conducted experiments for different σ_i^2 and $\boldsymbol{\mu}_i$ in generating the initial ensemble. It is found that the variances σ_i^2 do not affect much the final results, so in these experiments we keep $\sigma_i^2 = 9$ fixed, only allowing $\boldsymbol{\mu}$ to vary. Fig. 4 displays the results for the experiment with $\boldsymbol{\mu} = (9, 9, 9, 9)$. They correspond to those shown in the Fig. 3 of LK07b. Like the latter, most of the transfers are essentially zero. One of the nonzero transfers is $T_{3\to 2}$. See Fig. 4; also see Fig. 5 for a close-up. It is negative through the time, with a time average of -3.8. This is consistent with the $T_{3\to 2}$ in LK07b, which is positive (refer to the definition of relative entropy (12) for a relation between D and H). The difference is that here it is not a constant, but oscillates throughout; another difference is that it is far smaller in magnitude than its counterpart in LK07b.

The largest difference between the result here and that in LK07b is that there are two nonzero transfers here, and the dominant one is $T_{2\to 4}$, as shown in Fig. 4. In fact, $T_{3\to 2}$ is almost negligible in comparison with $T_{2\to 4}$, which averages to 20 and is nearly constant through the time. So it is $T_{2\to 4}$ that makes the counterpart of the $T_{3\to 2}$ in LK07b. Note that $T_{2\to 4}$ and $T_{3\to 2}$ stand for information flow in the opposite direction between the two modes—components (x_1, x_2) stand for the lower frequency mode in the truncation, while (x_3, x_4) for the higher mode (refer to LK07b for the derivation of (38)–(41)). In the present study, the computed result shows that information flows primarily from the lower mode to the higher mode, though weak information flow in the opposite direction has also been identified ($T_{3\to 2}$).

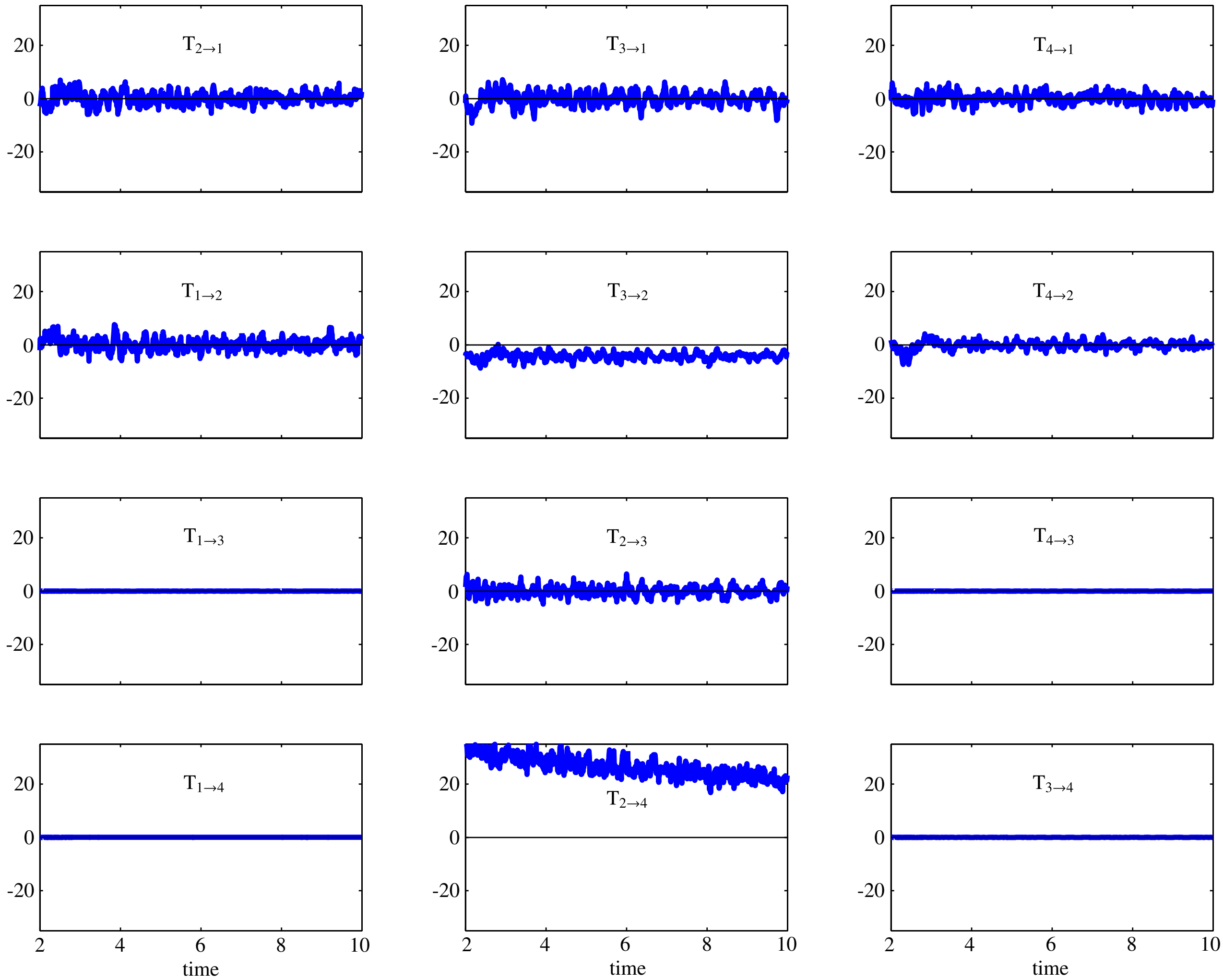


Fig. 4. Information transfers between different pairs of components with $\boldsymbol{\mu} = (9, 9, 9, 9)$. The series prior to $t = 2$ are not shown because some trajectories are still outside the computational domain by that time.

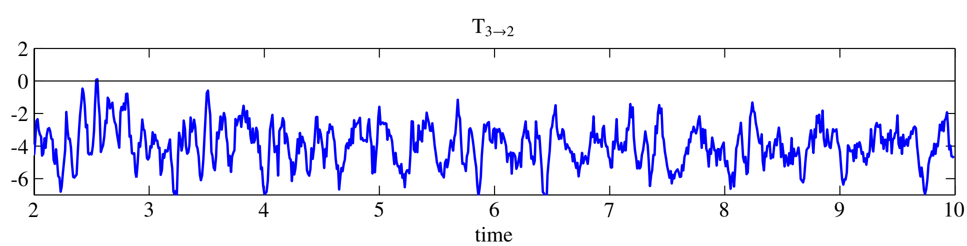


Fig. 5. A close-up of $T_{3\to 2}$ in Fig. 4.

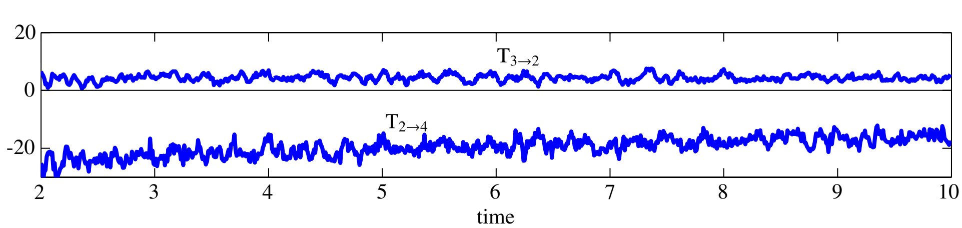


Fig. 6. Same as Fig. 4, but with $\boldsymbol{\mu} = (9, 9, -9, -9)$. Transfers other than $T_{2\to 4}$ and $T_{3\to 2}$ are not significantly different from zero.

The above result is very robust. Experiments with different $\boldsymbol{\mu}$ on $[-9, 9] \times [-9, 9] \times [-9, 9] \times [-9, 9]$ have been conducted, all yielding a $T_{2\to 4}$ large in magnitude, except that $\boldsymbol{\mu} = 0$ which makes all the transfers vanish. $T_{2\to 4}$ may be positive or negative, indicating that X_2 may increase or decrease the predictability of X_4. Plotted in Fig. 6 is the result of an experiment with $\boldsymbol{\mu} = (9, 9, -9, -9)$. The computed $T_{2\to 4}$ is approximately -20. In some of the experiments, discernible small $T_{3\to 2}$ as that in Fig. 5 may

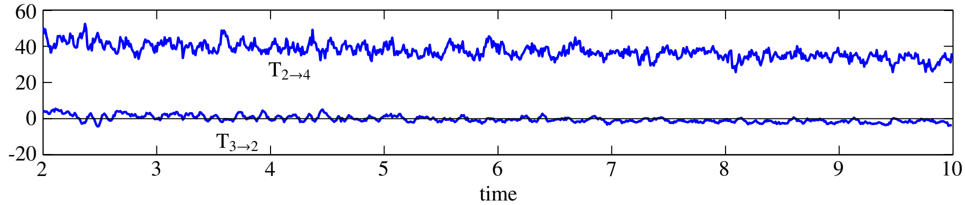


Fig. 7. Same as Fig. 4, but with $\sigma_2 = 4$. Transfers other than $T_{2\to4}$ and $T_{3\to2}$ are not significantly different from zero.

also be identified; in others it is not significantly different from zero. For example, in both Figs. 4(b) and 6 they are not zero (±4 on average), though very small in comparison to $T_{2\to4}$; By observation it is found that when $\mu_1 = \mu_2 < 0$, $T_{3\to2}$ vanishes, leaving $T_{2\to4}$ the only transfer. An interesting observation is that, if $T_{3\to2}$ is not zero, it always appears with a sign opposite to $T_{2\to4}$. Figs. 4 and 6 give two such examples. That is to say, if the lower frequency mode increases the predictability of the higher frequency mode, then the latter decreases the predictability of the former; and vice versa. This result is very interesting and has important implications for realistic problems, though the TBS is just a highly idealized model. We will give some discussion in the following section.

Experiments have also been conducted with varying initial variances, σ_i^2. Again, all of them result in a distinctly large $T_{2\to4}$. Shown in Fig. 7 is the result of an experiment, with $\boldsymbol{\mu}$ same as that in Fig. 4, but $(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2) = (9, 4, 9, 9)$. Since $\sigma_2^2 = 4$ is significantly smaller than others, initially we have a better knowledge of X_2, or have more information for X_2. It is thus expected that more information may be transferred from X_2 to X_4. In other words, $T_{2\to4}$ is expected to be much larger than that in Fig. 4. The time evolution of $T_{2\to4}$ displayed in Fig. 7 confirms this expectation.

8. Discussion and conclusions

Information transfer (or information flow as may be referred to) with respect to relative entropy was formulated to study how predictability varies locally as a dynamical system evolves. The resulting transfer measure (35) is in a form similar to that with respect to absolute entropy in previous studies [45], but for two terms modified in the formula. In the present context, a transfer from a component (say, X_2) to another component (say, X_1), written $T_{2\to1}$, tells the predictability change of X_1 due to X_2. When $T_{2\to1} < 0$, it means that the evolution of X_2 makes X_1 less predictable; on the other hand, a positive $T_{2\to1}$ implies that X_2 increases the predictability of X_1. The information transfer thus formulated has some nice properties. First of all, it satisfies the observed asymmetry requirement: the transfers between two components are asymmetric or directional. In particular, given a component, if its evolution is independent of another, then there is no information flowing from the latter, while in the same time the transfer in the opposite direction need not be zero. In other words, between two components, information transfer in one direction carries no implication of that in the other direction, in contrast to the transfer of other physical properties such as energy [54].

It should be pointed out that what the formalism yields is the *direct* information transfer/flow between two components. In a high-dimensional system, one can imagine that two components could be related, and hence indirect information transfers could take place, via a third party or more parties. As schematized in Fig. 8, $T_{2\to1} = 0$, i.e., no transfer exists from x_2 to x_1. However, if $T_{3\to1}$ and $T_{2\to3}$ are not zero, information does flow indirectly from x_2 to x_1 through x_3, adopting a route

$$x_2 \longrightarrow x_3 \longrightarrow x_1.$$

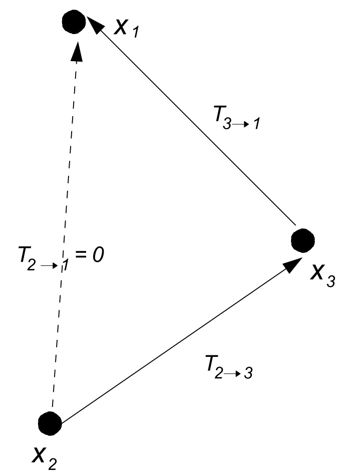


Fig. 8. The route of information transferring from x_2 to x_1 in the system (43)–(45). The transfer is through x_3; no direct transfer takes place.

A simple system that may mimic the schematic is

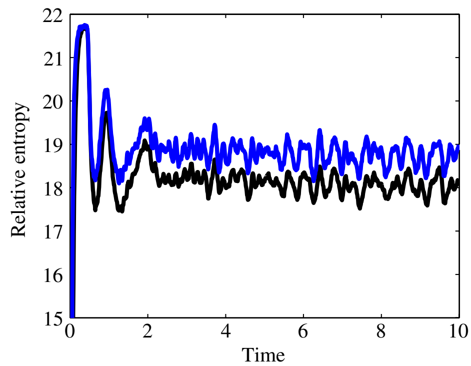
$$\frac{dx_1}{dt} = F_1(x_1, x_3), \tag{43}$$

$$\frac{dx_2}{dt} = F_2(x_1, x_2, x_3), \tag{44}$$

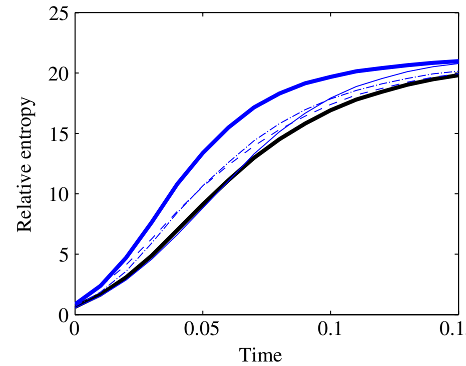
$$\frac{dx_3}{dt} = x_2. \tag{45}$$

We see that F_1 is independent of x_2, so by the theorem of causality, the transfer from x_2 to x_1 is zero: $T_{2\to1} = 0$. If F_1 and F_2 are such that $T_{2\to3}$ and $T_{3\to1}$ do not vanish, we get what Fig. 8 displays. For this particular system, one may argue that, with (45), x_3 can be expressed in terms of x_2 which, after substituted into (43) and (44), implies a direct transfer to x_1. This is not the case, as by substitution, the system is changed, with dimensionality reduced from three to two, so are the components. This is equivalent to saying that, in the schematic, the three-node system has been changed into a different, two-node system, and hence the information transfer must also be changed; the new x_2 essentially takes the composite effect of the original x_2 and x_3. The resulting information transfer is hence from both x_2 and x_3 in the original system and is, therefore, nonzero. From this example one can see, by computing the information transfer/flow as formulated, it is easy to identify the direct and indirect componentwise relations, plus the route through which the indirect relations are formed. This will be useful in the practical fields such as dependency identification among the nodes of a complex network.

The formalism was applied to the study of the information flow between the first two modes of a truncated Burgers system (TBS). The TBS is conservative, with energy and Hamiltonian being the first integrals (see Appendix D); it also conserves absolute entropy, as $\frac{dH}{dt} = E(\nabla \cdot \mathbf{F}) = 0$ with the $\mathbf{F}$ in (38)–(41) substituted. But its information in terms of relative entropy is not. In fact, substituting in (19) for the initial distribution q adopted in the experiments, one



(a) Evolution of predictability.



(b) Close-up.

Fig. 9. (a) Evolution of predictability with different choices of initial distributions. The lower thick line is the relative entropy with the parameters the same as for Fig. 4; the upper thick line is the same except that the initial variance for X_2 is smaller ($\sigma_2^2 = 4$). (b) A close-up of (a). The thin lines lying between the two bounds are the relative entropies with $\sigma_1^2 = 4$ (solid), $\sigma_3^2 = 4$ (dashed), and $\sigma_4^2 = 4$ (dot–dashed), respectively. (Other parameters are the same as for Fig. 4.)

obtains

$$\frac{dD}{dt} = \sum_{i=1}^{4} E\left(\frac{x_i - \mu_i}{\sigma_i^2} F_i\right). \tag{46}$$

With the F_i as defined in (38)–(41), it is easy to see that $\sum_i x_i F_i = 0$. So when σ_i^2 are all the same, say σ^2,

$$\frac{dD}{dt} = \sum_{i=1}^{4} \frac{\mu_i}{\sigma^2} E(F_i).$$

The terms on the right hand side are the second moments of the state variables; they generally do not sum to zero, unless $\mu_i = 0$. Therefore generally TBS does not conserve D or relative entropy.

Corresponding to the above fact are the computational results for the 12 transfers. Unless $\mu_i = 0$ in the case when σ_i^2 are identical, they do not add up to zero. But among them all are essentially zero, save for a strong transfer between the sine components from the low-frequency mode to the high-frequency mode ($T_{2\to4}$), plus a weak transfer from the high-frequency cosine component to the low-frequency sine component ($T_{3\to2}$). The latter is very small in comparison to the former, and may vanish in some cases. But, interestingly, if it does not vanish, it always carries a sign opposite to that of the former. In other words, if knowledge of the low-frequency mode increases the predictability of the high-frequency mode, the feedback, if any, is always to reduce the predictability of the low-frequency mode, and vice versa. As a whole, the information flow from the low-frequency mode is dominantly important; the flow in the opposite direction is generally negligible.

The asymmetry with the inter-modal information flow implies that, for this idealized TBS, the predictability of the high-frequency mode could be controlled by the low-frequency mode; specifically, the predictability of X_4 could be controlled by that of X_2. In other words, more precise knowledge of X_2 is expected to benefit the determination of X_4 in the future. Indeed this seems to be the case at the early stage of the marginal relative entropy evolution, i.e., the evolution of D_4 in our available experiments with different σ_2^2 (not shown). (But after a certain period, say after $t = 1$, the system loses memory about the initial condition because of its chaotic behavior.) Of course, here it is too early to reach a conclusion, as by time $t = 2$ many trajectories have not yet entered the computational domain for the sample space, making the computed ρ_4 hence D_4 for the early stage not reliable. There is still much work to be done along this line; we will defer this to later studies with more realistic systems.

The outstanding asymmetric transfer $T_{2\to4}$ may also help to improve the prediction of the system by adopting a better strategy of observing platform design. Specifically, it implies that precise observation of X_2 could be more important than that of other components of the system. We have rerun the experiment in Fig. 4, but with variance reduced to 4 for components 1, 2, 3, and 4, respectively. Accordingly the joint relative entropy D is computed and plotted. Note here D can be computed to substantial accuracy at all time instants with (46), where only moments of the random variables are involved: we just need to evaluate those moments by taking ensemble means, instead of coarse-graining the sample space. The computed Ds are plotted in Fig. 9. In Fig. 9(a), the two thick lines that form the lower and upper bounds are the D for the standard experiment (all σ_i^2 are 9), and the D for that with $\sigma_2^2 = 4$. Clearly the predictability is significantly improved by reducing σ_2^2 from 9 to 4. All other experiments result in evolution lines lying in between. For clarity, they are not displayed here, but displayed in a close-up plot (Fig. 9(b)) with thin lines. Obviously, for a better forecast of the system, precise observation of X_2 is more important than precise observation of other components. If, during the forecast, observational data are taken in, the result implies that X_2 should be identified as the primary variable for data assimilation.

The above causal relation between processes on different frequencies or scales is very important in that it has implications for one of the major problems in turbulence research and atmosphere–ocean science, i.e., the parameterization of unresolved or subgrid processes in numerical models. In a turbulent flow, there is a continuous spectrum of processes of all scales, but even with the most powerful computers to date it is impossible to resolve all the scales. The system is therefore not closed. The unresolved processes must be represented, or parameterized, with the resolved dynamics to fulfill the closure. The above result with the TBS implies that this kind of parameterization seems to work, as the predictability of the small-scale (high-frequency) mode is controlled by the large-scale (low-frequency) mode. Of course, one cannot draw conclusions from such a highly idealized system, although originally the TBS was introduced as a prototype of the atmosphere for the study of dynamical closure [55,56].

The importance of the inter-scale information transfer is not only out of the above practical concern; it is also an important physical problem in nature. For example, the North Atlantic Oscillation (NAO), the dominant mode controlling the wintertime climate of the North America and Europe, is believed to be driven by the synoptic eddies with a time scale of several days to weeks. How the NAO interacts with the eddies in the stormy boreal winters is a continuing challenging problem in atmospheric research because it is highly nonlinear in nature. Compounding the problem is that the interaction is essentially two way. That is to say, while the eddy-driven origin is an issue, the NAO also causes the growth and decay of the eddies. This work is expected to be useful in the investigation of these problems, particularly in the investigation of the causal relation in a quantitative way.

Another important problem concerns how uncertainty, and hence predictability, propagates in physical space. This is a problem naturally arising in fields such as material science, nanotechnology, and atmosphere–ocean science where ensemble prediction is used. This question actually may be posed for any problems governed by a partial differential equation (PDE). To illustrate how it may be approached, consider a Burgers equation. A dynamical system of large dimensionality can be formed by discretizing the space. The differencing may be fulfilled using a three point scheme. The resulting ODEs are thence connected to each other through the grid. Each ODE is tagged with the location in physical space, as well as a component in the system. The information transfer between the components then forms a flow of information in the space. Specifically, an ODE is connected to two other ODEs in the immediate neighborhood; it does not depend on other components. So for each point in the space, there are two transfers from ahead and back. By drawing these transfers, one obtains a flow forward and a flow backward, revealing how predictability changes due to uncertainty propagation. Theoretically this can be done, but practically it is hampered by the formidable computational job in evaluating the joint density in (35). For example, consider a five-dimensional joint density. Allowing 10 draws for each dimension, this totals 10^5 ensemble members. While ensembles of this size might be feasible for the Burgers equation, in dealing with realistic problems usually one can only handle ensembles of size in the order of hundreds or even tens. To overcome this difficulty, we need to simplify the rigorously derived formula (35), usually through problem-specific approximation. We leave this to future studies.

We remark that, when looking at the fractal measure of a strange attractor, one should use the measure-theoretic entropy namely the Kolmogorov–Sinai entropy (K–S entropy henceforth) [58,59], rather than the entropy in the usual sense, as the latter diverges. (That is the reason why in the TBS example only finite time evolution, not asymptotic behavior, is examined.) In this case, we need to establish a formalism of information transfer with respect to K–S entropy. Theoretically this should be feasible, as the system modified using the present strategy, i.e., after a component is frozen instantaneously as a parameter, is still a dynamical system in the usual sense, and hence its K–S entropy exists. The difficulty comes from the technical aspect, considering, for example, the supremum taken over all the partitions of the phase space. Besides, the K–S entropy is not really an entropy, but an entropy rate (entropy per unit time). The resulting transfer measure, if existing, might be quite different from what we have obtained, either with respect to Shannon entropy or with respect to relative entropy. How the present study may be generalized to have K–S entropy included is still a challenging research issue.

It should also be pointed out that, in many real problems such as the NAO mentioned above, the dynamics is not given explicitly, but in the form of observed data. In order for the formalism to apply, we first need to estimate from the series of observations (usually time series) the dynamical equations, then apply the result such as (35) to the estimated system. In a forthcoming paper, this will be treated within a more generic setting, with stochasticity taken into account. (The theoretical part is seen in [51].) The resulting transfer measure will be completely data based. That is to say, in that case one will be able to compute the transfer measure directly from the observations; no *a priori* knowledge of the dynamics is needed.

Acknowledgments

Richard Kleeman read through an early version of the manuscript: his insightful remarks are sincerely appreciated. The author is also grateful to three anonymous referees, whose comments helped to improve the presentation. This study was partially supported by the National Science Foundation of China (NSFC) under Grant No. 41276032 to NUIST, by Jiangsu Government through the "Jiangsu Specially-Appointed Professor Program" and "Priority Academic Program Development of Jiangsu Higher Education Institutions" (PAPD), and by the Ministry of Finance of China through the Basic Research Funding to China Institute for Advanced Study.

Appendix A. Proof of Theorem 1

Eq. (18) was originally obtained in LK05. It can be proved with the aid of the Liouville equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{F}\rho) = 0 \tag{A.1}$$

corresponding to the dynamical system (4) (cf. [53]). Multiplication of (A.1) by $-(1+\log\rho)$ gives

$$\begin{aligned}-\frac{\partial \rho \log \rho}{\partial t} &= \mathbf{F}\cdot\nabla(\rho\log\rho) + \rho(1+\log\rho)\nabla\cdot\mathbf{F} \\ &= [\log\rho\nabla\cdot(\mathbf{F}\rho) + \mathbf{F}\cdot\nabla\rho] + \rho\nabla\cdot\mathbf{F} \\ &= \nabla\cdot(\rho\log\rho\mathbf{F}) + \rho\nabla\cdot\mathbf{F}.\end{aligned}$$

Integrate over $\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ and notice the assumption of vanishing ρ at the boundaries. This results in

$$\frac{dH}{dt} = \int_\Omega \rho\nabla\cdot\mathbf{F}\,d\mathbf{x} = E(\nabla\cdot\mathbf{F}).$$

Eq. (19) is about the evolution of relative entropy D. By the definition (12) and using the Liouville equation, we have

$$\begin{aligned}\frac{dD}{dt} &= -\frac{dH}{dt} - \int_\Omega \frac{\partial\rho}{\partial t}\log q\,d\mathbf{x} \\ &= -E(\nabla\cdot\mathbf{F}) + \int_\Omega (\nabla\cdot\rho\mathbf{F})\log q\,d\mathbf{x} \\ &= -E(\nabla\cdot\mathbf{F}) - \int_\Omega \rho\mathbf{F}\cdot\nabla\log q\,d\mathbf{x} \\ &= -E(\nabla\cdot\mathbf{F}) - E\left(\mathbf{F}\cdot\nabla\log q\right),\end{aligned}$$

where integration by parts has been used.

To prove (20), first integrate (A.1) with respect to $(x_2, x_3, \ldots, x_n)$ over Ω_{2n} to get the evolution equation for ρ_1:

$$\frac{\partial\rho_1}{\partial t} + \int_{\Omega_{2n}} \frac{\partial(F_1\rho)}{\partial x_1}\,dx_2\ldots dx_n = 0. \tag{A.2}$$

Multiplying $-(1+\log\rho_1)$ and following the same procedure as above, one easily obtains (20).

By definition the marginal relative entropy of X_1 is

$$D_1 = -H_1 - \int_{\Omega_1} \rho_1\log q_1\,dx_1,$$

where q_1 is independent of time. Take derivatives on both sides with respect to t to get

$$\frac{dD_1}{dt} = -\frac{dH_1}{dt} - \int_{\Omega_1}\frac{\partial\rho_1}{\partial t}\log q_1\,dx_1.$$

Substitute (A.2) for $\frac{\partial\rho_1}{\partial t}$ and (21) follows. □

Appendix B. Proof of Proposition 3

This is a result obtained in LK07b; we rewrite the proof here for completeness.

We know from Section 4 that Φ and its components are invertible. $\Phi_{\not{2}}$ is hence also invertible, and its inverse is

$$\Phi_{\not{2}}^{-1}: \begin{cases} x_1 = y_1 - \Delta t \cdot F_1(y_1, x_2, y_3, \ldots, y_n) + O(\Delta t^2), \\ x_3 = y_3 - \Delta t \cdot F_3(y_1, x_2, y_3, \ldots, y_n) + O(\Delta t^2), \\ \vdots \quad \vdots \\ x_n = y_n - \Delta t \cdot F_n(y_1, x_2, y_3, \ldots, y_n) + O(\Delta t^2), \end{cases} \tag{B.1}$$

and

$$J_{\not{2}}^{-1} = \det\left[\frac{\partial(x_1, x_3, \ldots, x_n)}{\partial(y_1, y_3, \ldots, y_n)}\right]$$
$$= 1 - \Delta t \sum_{i \neq 2} \frac{\partial F_i}{\partial x_i} + O(\Delta t^2). \tag{B.2}$$

By (31),

$$\mathcal{P}_{\hat{2}}\rho(y_1, y_3, \ldots, y_n) = \rho\left(\Phi_{\not{2}}^{-1}(y_1, y_3, \ldots, y_n)\right)\left|J_{\not{2}}^{-1}\right|. \tag{B.3}$$

So

$$(\mathcal{P}_{\hat{2}}\rho)_1(y_1) = \int_{\Omega_{3n}} \rho_{\not{2}}(y_1 - \Delta t\, F_1, y_3 - \Delta t\, F_3, \ldots, y_n - \Delta t\, F_n)$$
$$\times \left|J_{\not{2}}^{-1}\right| dy_3 \ldots dy_n + O(\Delta t^2)$$
$$= \int_{\Omega_{3n}} \rho_{\not{2}}(\mathbf{y}_{\not{2}} - \Delta t\, \mathbf{F}_{\not{2}})\left|J_{\not{2}}^{-1}\right| dy_3 \ldots dy_n + O(\Delta t^2), \tag{B.4}$$

where $\mathbf{F}_{\not{2}} = (F_1, F_3, \ldots, F_n)$ are understood as functions of $(y_1, x_2, y_3, \ldots, y_n)$. Make a change of variables, $x_i = y_i - \Delta t \cdot F_i(y_1, x_2, y_3, \ldots, y_n)$, for $i = 3, 4, \ldots, n$. The Jacobian associated with this transformation is

$$J_{3n} = \det\left[\frac{\partial(y_3, y_4, \ldots, y_n)}{\partial(x_3, x_4, \ldots, x_n)}\right]$$
$$= 1 + \Delta t \sum_{i=3}^{n} \frac{\partial F_i}{\partial x_i} + O(\Delta t^2), \tag{B.5}$$

which gives

$$\left|J_{\not{2}}^{-1}\right| \cdot |J_{3n}| = 1 - \Delta t \frac{\partial F_1}{\partial x_1} + O(\Delta t^2). \tag{B.6}$$

With these substituted (B.4) becomes

$$(\mathcal{P}_{\hat{2}}\rho)_1(y_1) = \int_{\Omega_{3n}} \rho_{\not{2}}(y_1 - \Delta t\, F_1(y_1, x_2, x_3, \ldots, x_n))$$
$$\cdot \left|J_{\not{2}}^{-1}\right| \cdot |J_{3n}|\, dx_3 \ldots dx_n$$
$$= \int_{\Omega_{3n}} \rho_{\not{2}}(y_1 - \Delta t\, F_1(y_1, x_2, x_3, \ldots, x_n))$$
$$\cdot \left(1 - \Delta t \frac{\partial F_1}{\partial x_1}\right) dx_3 \ldots dx_n + O(\Delta t^2)$$
$$= \left[\rho_{\not{2}}(y_1, x_3, \ldots, x_n) - \Delta t \frac{\partial \rho_{\not{2}}}{\partial y_1} F_1\right]$$
$$\times \left(1 - \Delta t \frac{\partial F_1}{\partial x_1}\right) dx_3 \ldots dx_n + O(\Delta t^2)$$
$$= \rho_1(y_1) - \Delta t \cdot \int_{\Omega_{3n}} \left[\frac{\partial F_1}{\partial x_1} \rho_{\not{2}}(y_1, x_3, \ldots, x_n)\right.$$
$$\left. + F_1 \frac{\partial \rho_{\not{2}}(y_1, x_3, \ldots, x_n)}{\partial y_1}\right] dx_3 \ldots dx_n$$
$$+ O(\Delta t^2). \tag{B.7}$$

Since x_1 and y_1 are interchangeable up to an order of Δt, the two terms in the bracket can be combined, with the residual going to the higher order terms. That is to say,

$$(\mathcal{P}_{\hat{2}}\rho)_1(y_1) = \rho_1(y_1) - \Delta t \cdot \int_{\Omega_{3n}} \frac{\partial F_1 \rho_{\not{2}}}{\partial y_1} dx_3 \ldots dx_n$$
$$+ O(\Delta t^2). \quad \square$$

Appendix C. Proof of Theorem 2

Subtract $D_1(t)$ from (33) to get

$$\Delta D_{1\not{2}} = D_{1\not{2}}(t + \Delta t) - D_1(t)$$
$$= -\Delta H_{1\not{2}} - \int_{\Omega} (\mathcal{P}_{\hat{2}}\rho)_1(y_1) \log q_1(y_1) \cdot \rho(x_2|x_1, x_3, \ldots, x_n)$$
$$\cdot \rho_{3\ldots n}(x_3, \ldots, x_n)\, dy_1 dx_2 \ldots dx_n$$
$$+ \int_{\Omega} \rho_1(x_1) \log q_1(x_1)\, dx_1$$
$$\equiv -\Delta H_{1\not{2}} - D_* + \int_{\Omega} \rho_1 \log q_1\, dx_1. \tag{C.1}$$

In this equation, $\Delta H_{1\not{2}}$ has already been obtained in LK07b. We need to compute D_*. Note x_1 and y_1 coexist in the expression, the latter being $x_1 + F_1 \Delta t$. Perform a Taylor series expansion around $(y_1, x_2, x_3, \ldots, x_n)$ to get rid of x_1:

$$\rho(x_2|x_1, x_3, \ldots, x_n) = \frac{\rho}{\rho_{\not{2}}}(x_1, x_2, \ldots, x_n)$$
$$= \frac{\rho}{\rho_{\not{2}}} + \frac{\partial \frac{\rho}{\rho_{\not{2}}}}{\partial y_1} \cdot (-F_1 \Delta t) + O(\Delta t^2)$$
$$= \rho(x_2|y_1, x_3, \ldots, x_n) + \frac{\rho}{\rho_{\not{2}}^{\,2}} \frac{\partial \rho_{\not{2}}}{\partial y_1} F_1 \Delta t$$
$$- \frac{1}{\rho_{\not{2}}} \frac{\partial \rho}{\partial y_1} F_1 \Delta t + O(\Delta t^2)$$
$$= \rho(x_2|y_1, x_3, \ldots, x_n)$$
$$+ \rho(x_2|y_1, x_3, \ldots, x_n) \cdot \frac{\partial \log \rho_{\not{2}}}{\partial y_1} F_1 \Delta t$$
$$- \frac{1}{\rho_{\not{2}}} \frac{\partial \rho}{\partial y_1} F_1 \Delta t + O(\Delta t^2), \tag{C.2}$$

where the variables without independent variables explicitly written out are tacitly supposed to be functions of $(y_1, x_2, x_3, \ldots, x_n)$. With this expansion and (34) from Proposition 3,

$$D_* = \int_{\Omega} \log q_1(y_1) \cdot \left(\rho_1(y_1) - \Delta t \int_{\Omega_{3n}} \frac{\partial F_1 \rho_{\not{2}}}{\partial y_1} dx_3 \ldots dx_n\right)$$
$$\cdot \left(\rho(x_2|y_1, x_3, \ldots, x_n) + \rho(x_2|y_1, x_3, \ldots, x_n)\right.$$
$$\left. \cdot \frac{\partial \log \rho_{\not{2}}}{\partial y_1} F_1 \Delta t - \frac{1}{\rho_{\not{2}}} \frac{\partial \rho}{\partial y_1} F_1 \Delta t\right)$$
$$\cdot \rho_{3\ldots n}(x_3, \ldots, x_n) dy_1 dx_2 \ldots dx_n + O(\Delta t^2)$$
$$= \int_{\Omega} \log q_1(y_1) \cdot \rho_1(y_1)$$
$$\cdot \rho(x_2|y_1, x_3, \ldots, x_n)\, dy_1 dx_2 \ldots dx_n$$
$$+ \Delta t \int_{\Omega} \log q_1(y_1) \cdot \rho_1(y_1)$$
$$\cdot \left[\rho(x_2|y_1, x_3, \ldots, x_n) \frac{\partial \log \rho_{\not{2}}}{\partial y_1} F_1 - \frac{1}{\rho_{\not{2}}} \frac{\partial \rho}{\partial y_1} F_1\right]$$

$$
\begin{aligned}
&\cdot \rho_{3\ldots n}(x_3,\ldots,x_n)\,dy_1 dx_2\ldots dx_n \\
&- \Delta t \int_\Omega \log q_1(y_1)\cdot\left(\int_{\Omega_{3n}} \frac{\partial F_1\rho_{\not 2}}{\partial y_1} dx_3\ldots dx_n\right) \\
&\cdot \rho(x_2|y_1,x_3,\ldots,x_n) \\
&\cdot \rho_{3\ldots n}(x_3,\ldots,x_n)\,dy_1dx_2\ldots dx_n + O(\Delta t^2) \\
\equiv\ & (I)+(II)+(III)+O(\Delta t^2).
\end{aligned}
$$

Now evaluate the three terms one by one. For convenience, all the y_1 are replaced by x_1. This is legitimate as now y_1 is a dummy variable. The first term is

$$
\begin{aligned}
(I) &= \int_\Omega \log q_1\cdot\rho_1\cdot\frac{\rho}{\rho_{\not 2}}\cdot\rho_{\not 1\not 2}\,d\mathbf{x} \\
&= \int_{\Omega_1}\rho_1(x_1)\log q_1(x_1)\,dx_1.
\end{aligned}
\tag{C.3}
$$

At the second step, we first integrate $\frac{\rho(\mathbf{x})}{\rho_{\not 2}}$ with respect to x_2 (all other parts are independent of x_2) to get 1, then take the integral with respect to $x_3,\ldots,x_n$ and eliminate $\rho_{\not 1\not 2}$.

For the second part,

$$
\begin{aligned}
(II) &= \Delta t\int_\Omega \log q_1(x_1)\cdot\rho_1(x_1) \\
&\quad\times\left[\frac{\rho}{\rho_{\not 2}}\frac{\partial\log\rho_{\not 2}}{\partial x_1}F_1 - \frac{1}{\rho_{\not 2}}\frac{\partial\rho}{\partial x_1}F_1\right]\rho_{\not 1\not 2}\,d\mathbf{x} \\
&= -\Delta t\int_\Omega \rho_1\log q_1\cdot\frac{\rho_{\not 2}}{\rho}\cdot\frac{\partial\rho/\rho_{\not 2}}{\partial x_1}\cdot F_1\cdot\frac{\rho\rho_{\not 1\not 2}}{\rho_{\not 2}}\,d\mathbf{x} \\
&= -\Delta t\int_\Omega F_1\rho_1\log q_1\cdot\frac{\partial}{\partial x_1}\left(\frac{\rho_{\not 1\not 2}\,\rho}{\rho_{\not 2}}\right)d\mathbf{x}.
\end{aligned}
\tag{C.4}
$$

Using the notations (24) and (25):

$$
\theta_{2|1} = \frac{\rho_{\not 1\not 2}\,\rho}{\rho_{\not 2}},
$$

$$
\Theta_{2|1} = \int_{\Omega_{3n}}\theta_{2|1}(\mathbf{x})\,dx_3\ldots dx_n
$$

simplifies the above formula to be

$$
\begin{aligned}
(II) &= -\Delta t\int_\Omega F_1\rho_1\log q_1\cdot\frac{\partial\theta_{2|1}}{\partial x_1}\,d\mathbf{x} \\
&= \Delta t\int_\Omega \frac{\partial(F_1\rho_1\log q_1)}{\partial x_1}\cdot\theta_{2|1}\,d\mathbf{x}.
\end{aligned}
\tag{C.5}
$$

The third term (III) may be equally simplified,

$$
\begin{aligned}
(III) &= -\Delta t\int_\Omega\left[\log q_1(x_1)\cdot\left(\int_{\Omega_{3n}}\frac{\partial F_1\rho_{\not 2}}{\partial x_1}dx_3\ldots dx_n\right)\right] \\
&\quad\cdot\theta_{2|1}(\mathbf{x})\,d\mathbf{x}.
\end{aligned}
$$

The part in the square brackets is independent of $(x_3,\ldots,x_n)$. So integration can be performed on $\theta_{2|1}$ with respect to $(x_3,\ldots,x_n)$, which gives

$$
\begin{aligned}
(III) &= -\Delta t\int_{\Omega_1\times\Omega_2}\log q_1(x_1)\cdot\left(\int_{\Omega_{3n}}\frac{\partial F_1\rho_{\not 2}}{\partial x_1}dx_3\ldots dx_n\right) \\
&\quad\cdot\Theta_{2|1}(x_1,x_2)\,dx_1dx_2 \\
&= -\Delta t\int_\Omega \log q_1\cdot\frac{\partial F_1\rho_{\not 2}}{\partial x_1}\cdot\Theta_{2|1}\,d\mathbf{x}.
\end{aligned}
\tag{C.6}
$$

Combining (I), (II), and (III), one has

$$
\begin{aligned}
D_* &= \int_{\Omega_1}\rho_1\log q_1\,dx_1 + \Delta t\int_\Omega\frac{\partial(F_1\rho_1\log q_1)}{\partial x_1}\theta_{2|1}\,d\mathbf{x} \\
&\quad - \Delta t\int_\Omega \log q_1\frac{\partial F_1\rho_{\not 2}}{\partial x_1}\Theta_{2|1}\,d\mathbf{x} + O(\Delta t^2).
\end{aligned}
$$

So

$$
\begin{aligned}
\Delta D_{1\not 2} &= -\Delta H_{1\not 2} - D_* + \int_\Omega \rho_1\log q_1\,dx_1 \\
&= -\Delta H_{1\not 2} - \Delta t\int_\Omega\frac{\partial F_1\rho_1\log q_1}{\partial x_1}\theta_{2|1}\,d\mathbf{x} \\
&\quad + \Delta t\int_\Omega \log q_1\frac{\partial F_1\rho_{1\not 2}}{\partial x_1}\Theta_{2|1}\,d\mathbf{x} + O(\Delta t^2).
\end{aligned}
$$

Letting $\Delta t\to 0$, it becomes

$$
\begin{aligned}
\frac{dD_{1\not 2}}{dt} &= -\frac{dH_{1\not 2}}{dt} - \int_\Omega\frac{\partial F_1\rho_1\log q_1}{\partial x_1}\theta_{2|1}\,d\mathbf{x} \\
&\quad + \int_\Omega \log q_1\frac{\partial F_1\rho_{1\not 2}}{\partial x_1}\Theta_{2|1}\,d\mathbf{x}.
\end{aligned}
\tag{C.7}
$$

By Eq. (48) of LK07b,

$$
\begin{aligned}
\frac{dH_{1\not 2}}{dt} &= \int_\Omega(1+\log\rho_1)\frac{\partial F_1\rho_{\not 2}}{\partial x_1}\Theta_{2|1}\,d\mathbf{x} \\
&\quad + \int_\Omega F_1\rho_1\log\rho_1\frac{\partial\rho/\rho_{\not 2}}{\partial x_1}\rho_{\not 1\not 2}\,d\mathbf{x} \\
&= \int_\Omega(1+\log\rho_1)\frac{\partial F_1\rho_{\not 2}}{\partial x_1}\Theta_{2|1}\,d\mathbf{x} \\
&\quad + \int_\Omega F_1\rho_1\log\rho_1\frac{\partial\theta_{2|1}}{\partial x_1}\,d\mathbf{x} \\
&= \int_\Omega(1+\log\rho_1)\frac{\partial F_1\rho_{\not 2}}{\partial x_1}\Theta_{2|1}\,d\mathbf{x} \\
&\quad - \int_\Omega\frac{\partial(F_1\rho_1\log\rho_1)}{\partial x_1}\theta_{2|1}\,d\mathbf{x}.
\end{aligned}
\tag{C.8}
$$

In arriving at the last step, integration by parts has been performed together with the assumption of vanishing boundary density. Substituting (C.8) back to (C.7), one finally arrives at (23). □

Appendix D. Solution of (38)–(41)

It has been shown in [57] that the truncated Burgers system conserves momentum, energy, and Hamiltonian. The momentum conservation just gives the evolution equations themselves; the remaining two invariants, i.e., energy and Hamiltonian, can be utilized to have the 4D system integrated out. Hereafter we show how. But before moving on, a transformation is needed for the state variables:

$$
(q_1,p_1) = (x_1,x_2) \tag{D.1}
$$

$$
(q_2,p_2) = \left(\frac{1}{\sqrt{2}}x_3,\ \frac{1}{\sqrt{2}}x_4\right). \tag{D.2}
$$

In terms of (q_i,p_i), $i=1,2$, the original system (38)–(41) can be written in a canonical form:

$$
\frac{dq_i}{dt} = \frac{\partial\mathcal{H}}{\partial p_i}, \tag{D.3}
$$

$$
\frac{dp_i}{dt} = -\frac{\partial\mathcal{H}}{\partial q_i}, \quad (i=1,2) \tag{D.4}
$$

with q_i and p_i the configuration and momentum coordinates, respectively. In the equations, the Hamiltonian is

$$
\mathcal{H} = \frac{1}{\sqrt{2}}q_2(q_1^2-p_1^2) + \sqrt{2}q_1p_1p_2. \tag{D.5}
$$

By algebraic manipulation one sees that $d\mathcal{H}/dt = 0$; that is to say, $\mathcal{H}$ is a first integral. Another first integral is energy

$$
\mathcal{E} = \sum_{i=1}^4 x_i^2 = q_1^2+p_1^2+2\left(q_2^2+p_2^2\right). \tag{D.6}
$$

Indeed, it is easy to check that $d\mathcal{E}/dt = 0$. Moreover, $\mathcal{H}$ and $\mathcal{E}$ are in involution, i.e., the Poisson bracket of $\mathcal{H}$ and $\mathcal{E}$ vanishes:

$$\begin{aligned}[\mathcal{H}, \mathcal{E}] &= \frac{\partial \mathcal{H}}{\partial q_1}\frac{\partial \mathcal{E}}{\partial p_1} - \frac{\partial \mathcal{H}}{\partial p_1}\frac{\partial \mathcal{E}}{\partial q_1} + \frac{\partial \mathcal{H}}{\partial q_2}\frac{\partial \mathcal{E}}{\partial p_2} - \frac{\partial \mathcal{H}}{\partial p_2}\frac{\partial \mathcal{E}}{\partial q_2} \\ &= \left(\sqrt{2}q_2q_1 + \sqrt{2}p_1p_2\right)\cdot(2p_1) \\ &\quad - \left(-\sqrt{2}q_2p_1 + \sqrt{2}q_1p_2\right)\cdot(2q_1) \\ &\quad + \frac{1}{\sqrt{2}}\left(q_1^2 - p_1^2\right)\cdot(4p_2) - 4\sqrt{2}q_1p_1q_2 \\ &= 0. \end{aligned} \tag{D.7}$$

This property makes it possible to make a canonical transformation (e.g., [4,60]):

$$\bar{q}_i = \bar{q}_i(q_1, q_2; p_1, p_2) \tag{D.8}$$

$$\bar{I}_i = \bar{I}_i(q_1, q_2; p_1, p_2) \tag{D.9}$$

with the new momentum coordinates being the invariants $I_1 = \mathcal{E}$ and $I_2 = \mathcal{H}$. Let the generating function [4] of the second type be

$$G(q_1, q_2; I_1, I_2) = \int_{(q_1^0, q_1^0)}^{(q_1, q_2)} \sum_{i=1}^{2} p_i(q_1, q_2; I_1, I_2)\, dq_i, \tag{D.10}$$

where p_i are inverted from (D.5) and (D.6) such that they are expressed in terms of (q_1, q_2, I_1, I_2), then

$$p_i = \frac{\partial G}{\partial q_i}, \tag{D.11}$$

$$\bar{q}_i = \frac{\partial G}{\partial I_i}, \tag{D.12}$$

$$\bar{\mathcal{H}} = \mathcal{H} + \frac{\partial G}{\partial t} = \mathcal{H}, \tag{D.13}$$

and, in the new coordinates,

$$\frac{dI_i}{dt} = -\frac{\partial \bar{\mathcal{H}}}{\partial \bar{q}_i} = -\frac{\partial \mathcal{H}}{\partial \bar{q}_i}, \tag{D.14}$$

$$\frac{d\bar{q}_i}{dt} = \frac{\partial \bar{\mathcal{H}}}{\partial I_i} = \frac{\partial \mathcal{H}}{\partial I_i}. \tag{D.15}$$

In (D.14), $(I_1, I_2) = (\mathcal{E}, \mathcal{H})$ are constants, so $\partial \mathcal{H}/\partial \bar{q}_i = 0$; in other words, $\mathcal{H}$ is independent of $\bar{q}_i$, or is a function of I_i only. Let $\partial \mathcal{H}/\partial I_i \equiv \pi_i$ (it is easy to know that $\pi_2 = 1$). Eq. (D.15) then can be integrated out, and accordingly the new configuration coordinates are obtained:

$$\bar{q}_i = \pi_i t + c_i, \quad i = 1, 2 \tag{D.16}$$

where c_i are the integral constants that can determined by the initial conditions. Note $\bar{q}_i = \bar{q}_i(q_1, q_2; I_1, I_2)$ are known from (D.12), so (D.16) are two equalities about the variables $(q_1, q_2; I_1, I_2)$ given a fixed time t. They together with the first integrals $I_1 = \mathcal{E}$ [Eq. (D.6)] and $I_2 = \mathcal{H}$ [Eq. (D.5)] constitute the four constraints that determine a trajectory in the 4D phase space (q_1, q_2, p_1, p_2) (and hence the space (x_1, x_2, x_3, x_4)). An explicit expression for the phase portrait is very complicated and not helpful in this case. The reader is referred to Fig. 2 for such a trajectory.

References

[1] E.N. Lorenz, Deterministic non-periodic flow, J. Atmos. Sci. 20 (1963) 130–141.
[2] T. Leiber, On the impact of deterministic chaos on modern science and philosophy of science, Phil. & Tech. 4 (2) (1998) 23–50.
[3] R. Doerner, B. Hbinger, W. Martienssen, S. Grossmann, S. Thomas, Stable manifolds and predictability of dynamical systems, Chaos Solitons Fractals 10 (1999) 1759–1782.
[4] A.J. Lichtenberger, M.A. Lieberman, Regular and Chaotic Dynamics, Springer, New York, 1992.
[5] G.F. Carnevale, G. Holloway, Information decay and the predictability of turbulent flows, J. Fluid Mech. 116 (1982) 115–121.
[6] T. DelSole, Predictability and information theory. Part I: measures of predictability, J. Atmos. Sci. 61 (20) (2004) 2425.
[7] R. Ding, J.P. Li, Nonlinear finite-time Lyapunov exponent and predictability, Phys. Lett. A 364 (2007) 396–400.
[8] B.F. Farrell, Small error dynamics and predictability of atmospheric flows, J. Atmos. Sci. 47 (1990) 2409–2416.
[9] E. Kalnay, Atmospheric Modeling, Data Assimilation, and Predictability, Cambridge Univ Press, Cambridge, UK, 2003.
[10] R. Kleeman, Measuring dynamical prediction utility using relative entropy, J. Atmos. Sci. 59 (2002) 2057–2072.
[11] R. Kleeman, A. Majda, I. Timofeyev, Quantifying predictability in a model with statistical features of the atmosphere, Proc. Natl. Acad. Sci. 99 (2002) 15291–15296.
[12] R. Kleeman, A.J. Majda, Predictability in a model of geostrophic turbulence, J. Atmos. Sci. 62 (2005) 2864–2879.
[13] R. Kleeman, Statistical predictability in the atmosphere and other dynamical systems, Physica D 230 (2007) 65–71.
[14] R. Kleeman, Limits, variability and general behavior of statistical predictability of the mid-latitude atmosphere, J. Atmos. Sci. 65 (2008) 263–275.
[15] M. Mu, W. Duan, J. Wang, The predictability problems in numerical weather and climate prediction, Adv. Atmos. Sci. 19 (2002) 191–204.
[16] T. Schneider, S.M. Griffies, A conceptual framework for predictability studies, J. Clim. 12 (1999) 3133–3155.
[17] J. Shukla, Predictability in the midst of chaos: a scientific basis for climate forecasting, Science 282 (1998) 728–731.
[18] L.A. Smith, C. Ziehmann, K. Fraedrich, Uncertainty dynamics and predictability in chaotic systems, Q. J. R. Meteorol. Soc. 125 (1999) 2855–2886.
[19] M.K. Tippett, P. Chang, Some theoretical considerations on predictability of linear stochastic dynamics, Tellus A 55 (2) (2003) 148–157.
[20] J.J. Tribbia, D.P. Baumhefner, Scale interactions and atmospheric predictability: an updated perspective, Mon. Weather Rev. 132 (2004) 703–713.
[21] W. Duan, M. Mu, B. Wang, Conditional nonlinear optimal perturbations as the optimal precursorrs for El Nino-Southern Oscillation events, J. Geophys. Res. D 23105 (2004) 1029–1041.
[22] M. Ehrendorfer, J.J. Tribbia, Optimal prediction of forecast error covariances through singular vectors, J. Atmos. Sci. 54 (1997) 286–313.
[23] E.S. Epstein, Stochastic dynamic prediction, Tellus 21 (1969) 739–759.
[24] A.D. Kirwan Jr., M. Toner, L. Kantha, Predictability, uncertainty, and hyperbolicity in the ocean, Internat. J. Engrg. Sci. 41 (2003) 249–258.
[25] R. Kleeman, Information flow in ensemble weather predictions, J. Atmos. Sci. 64 (3) (2007) 1005–1016.
[26] C.E. Leith, Theoretical skill of Monte Carlo forecasts, Mon. Weather Rev. 102 (1974) 409–418.
[27] P.F.J. Lermusiaux, Uncertainty estimation and prediction for interdisciplinary ocean dynamics, in: J. Glimm, G. Karniadakis (Eds.), J. Comput. Phys. 217 (2006), 176–199 (Special issue of on "uncertainty quantification").
[28] R.N. Miller, L.L. Ehret, Ensemble generation for models of multimodal systems, Mon. Weather Rev. 130 (2002) 2313–2333.
[29] A.M. Moore, The dynamics of error growth and predictability in a model of the Gulf Stream. II: ensemble prediction, J. Phys. Oceanogr. 29 (1999) 762–778.
[30] T.N. Palmer, Predicting uncertainty in forecasts of weather and climate, Rep. Progr. Phys. 63 (2000) 71–116.
[31] Z. Toth, E. Kalnay, Operational ensemble prediction at the National Meteorological Center: practical aspects, Bull. Amer. Meteor. Soc. 74 (1993) 2317–2330.
[32] T.N. Palmer, Medium and extended range predictability and stability of the Pacific/North American mode, Q. J. R. Meteorol. Soc. 114 (1988) 691–713.
[33] J.J. Tribbia, Waves, information and local predictability, in: Workshop on Mathematical Issues and Challenges in Data Assimilation for Geophysical Systems: Interdisciplinary Perspectives, UCLA, February 22-25, IPAM, 2005.
[34] H.D.I. Abarbanel, N. Masuda, M.I. Rabinovich, E. Turner, Distribution of mutual information, Phys. Lett. A 281 (2001) 368–373.
[35] Arnhold, et al., A robust method for detecting interdependences: application to intracranially recorded EEG, Physica D 134 (1999) 419–430.
[36] A. Kaiser, T. Schreiber, Information transfer in continuous processes, Physica D 166 (2002) 43–62.
[37] K. Kaneko, Lyapunov analysis and information flow in coupled map lattices, Physica D 23 (1986) 436–447.
[38] M.G. Rosenblum, A.S. Pikovsky, J. Kurths, Phase synchronization of chaotic oscillators, Phys. Rev. Lett. 76 (1996) 1804–1807.
[39] K. Schindler-Hlavackova, M. Palus, M. Vejmelka, J. Bhattacharya, Causality detection based on information-theoretic approaches in time series analysis, Phys. Rep. 441 (2007) 1–46.
[40] T. Schreiber, Measuring information transfer, Phys. Rev. Lett. 85 (2) (2000) 461.
[41] J.A. Vastano, H.L. Swinney, Information transport in spatiotemporal systems, Phys. Rev. Lett. 60 (1988) 1773–1776.
[42] R. Metzler, Y. Bar-Yam, M. Kardar, Information flow through a chaotic channel: prediction and postdiction at finite resolution, Phys. Rev. E 70 (2004) 026250.

[43] X.S. Liang, R. Kleeman, Information transfer between dynamical system components, Phys. Rev. Lett. 95 (24) (2005) 244101.
[44] X.S. Liang, R. Kleeman, A rigorous formalism of information transfer between dynamical system components. I. Discrete mapping, Physica D 231 (2007) 1–9.
[45] X.S. Liang, R. Kleeman, A rigorous formalism of information transfer between dynamical system components. II. Continuous flow, Physica D 227 (2007) 173–182.
[46] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[47] P. Ao, Emerging of stochastic dynamical equalities and steady state thermodynamics from Darwinian dynamics, Commun. Theor. Phys. 49 (2008) 1073–1090.
[48] K. Sangil, R.M. Samelson, C. Snyder, Ensemble-based estimates of the predictability of wind-driven coastal ocean flow over topography, Mon. Wea. Rev. 137 (8) (2009) 2515–2537.
[49] Y. Tang, Z. Deng, X. Zhou, Y. Cheng, D. Chen, Interdecadal variation of ENSO predictability in multiple models, J. Clim. 21 (2008) 4811–4832.
[50] A.J. Majda, J. Harlim, Information flow between subspaces of complex dynamical systems, Proc. Natl. Acad. Sci. 104 (23) (2007) 9558–9563.
[51] X.S. Liang, Information flow within stochastic dynamical systems, Phys. Rev. E 78 (2008) 031113.
[52] E. Pereda, R.Q. Quiroga, J. Bhattacharya, Nonlinear multivariate analysis of neurophysiological signals, Prog. Neurobiol. 77 (1–2) (2005) 1–37.
[53] A. Lasota, M.C. Mackey, Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics, Springer, New York, 1994.
[54] X.S. Liang, A.R. Robinson, Localized multiscale energy and vorticity analysis II. Finite-amplitude instability theory and validation, Dyn. Atmos. Oceans. 44 (2007) 51–76.
[55] A.J. Majda, I. Timofeyev, Remarkable statistical behavior for truncated burgers-Hopf dynamics, Proc. Natl. Acad. Sci. USA 97 (23) (2000) 12413–12417.
[56] A.J. Majda, I. Timofeyev, Statistical mechanics for truncations of the Burgers–Hopf equation. A model for intrinsic stochastic behavior with scaling, Milan J. Math. 70 (1) (2002) 39–96.
[57] R.V. Abramov, G. Kovacic, A.J. Majda, Hamiltonian structure and statistically relevant conserved quantities for the truncated Burgers–Hopf equation, Comm. Pure Appl. Math. LVI (2003) 1–46.
[58] A.N. Kolmogorov, Entropy per unit time as a metric invariant automorphism, Dokl. Russ. Acad. Sci. 124 (1959) 754–755.
[59] Ya.G. Sinai, On the notion of entropy of a dynamical system, Dokl. Russ. Acad. Sci. 124 (1959) 768–771.
[60] S.N. Rasband, Dynamics, John Wiley & Sons, New York, 1983.