IOWA STATE UNIVERSITY
Digital Repository

2018

Information transfer in dynamical systems

Subhrajit Sinha
Iowa State University

Information transfer in dynamical systems

by

Subhrajit Sinha

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Program of Study Committee:
Umesh Vaidya, Major Professor
Nicola Elia
Venkataramana Ajjarapu
Baskar Ganapathysubramanian
Soumik Sarkar

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

DEDICATION

I would like to dedicate my thesis to all the wonderful teachers I have had over the years, without whom I would not have learned to ask questions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

xiv

ABSTRACT

Causality analysis has been a topic of research from the days of Aristotle. However, there has not been an universal definition of causality, with different researchers providing different definitions and measures of causality. In this work, we provide a new definition of causality in a dynamical system setting. In particular, we quantify how a state (or subspace) of a dynamical system influence any other state (or subspace). The quantification is in terms of what we call information transfer. Intuitively, information transfer quantifies how the entropy of a state (say y) changes due to evolution of any other state (say x) and this characterizes the causal structure and influence in a dynamical system. We show that the information transfer measure satisfies intuitions of causality and influence. In particular, we show our information transfer measure satisfy *(a) zero influence, (b) transfer asymmetry* and *(c) information conservation*. The one step information transfer is generalized to define n-step information transfer and this enables us to clearly distinguish between direct and indirect influence. With this, we show how in a dynamical system setting, when previous measures of causality fail to capture the true causal structure, our information transfer measure do capture the true causal structure in a dynamical system. Apart from identification of causal structure, we use the information transfer measure to characterize influence and demonstrate how this can be used to characterize stability in a power network. In particular, we show how to identify the states and generators which are responsible for instability of a power network. Moreover, we show how the connection between stability and information transfer can be used to predict phase transitions in complex systems. We also provide two algorithms to compute information transfer from time series data and show how this can be used for topology identification of linear networks and stability analysis of power networks. Finally, as a separate application, we characterize influence in a stock market and also show how the information transfer measure can be used to predict stock market crashes.

CHAPTER 1.   INTRODUCTION

Causality and influence characterization is a problem of utmost importance and finds its relevance in many different applications like world wide web and social media, biological networks, neural science, economics, finance etc. Usually, concepts of information theory are used in such applications and a study of the information flow between the components of the network throws light on causality and the influential nodes of the network. In (1) the authors use information based metric to characterize the most influential nodes in social networks. In neuroscience, concepts of information theory are used to understand how information flow in different parts of the brain (2) and identifying influence in gene regulatory networks (3; 4) .

Again, Clive Granger, in 1969 (5), provided his own definition of causality to study causality in economic models. This definition of causality, which is now known as Granger causality, is used in economic and financial networks to infer causal interactions from the time series data (6; 7; 8).

Causality detection in system theoretic setting was addressed much later in 2000, when Thomas Schreiber (9) defined transfer entropy between two states in a dynamical system. In 2005, Liang and Kleeman defined information transfer between dynamical system components (10) and provided explicit expressions for the information transfer between the states in a continuous time dynamical system.

1.1   Previous Work

Causality is the agency or efficacy that connects one process (the cause) with another (the effect), where the first is understood to be partly responsible for the second, and the second is dependent on the first. In general, a process has many causes, which are said to be causal factors for it, and all lie in its past. An effect can, in turn, be a cause of many other effects, which all lie in its future. Interpreting

causation as a deterministic relation means that if A causes B, then A must always be followed by B. In this sense, war does not cause deaths, nor does smoking cause cancer or emphysema. As a result, a probabilistic notion of causation is often used. Informally, A ("The person is a smoker") probabilistically causes B ("The person has now or will have cancer at some time in the future"), if the information that A occurred increases the likelihood of B's occurrence. This has led the philosophers and researchers to define causality in a probabilistic setting.

However, there has been no universally accepted definition of causality. One of the most commonly used notions of causality is Granger causality and was proposed by Clive Granger for studying causality in the economics framework and is normally tested in context of linear regression models. For example, consider a bivariate linear autoregressive model of two variables x_1 and x_2

$$x_1(t) = \sum_{j=1}^{p} A_{11,j} x_1(t-j) + \sum_{j=1}^{p} A_{12,j} x_2(t-j) + E_1(t)$$
$$x_2(t) = \sum_{j=1}^{p} A_{21,j} x_1(t-j) + \sum_{j=1}^{p} A_{22,j} x_2(t-j) + E_1(t)$$

where p is the maximum number of lagged observations included in the model (the model order), the matrix A contains the coefficients of the model and $E_1(t)$ and $E_2(t)$ are the prediction errors for each time series. If the variance of E_1(or E_2) is reduced by the inclusion of x_2(or x_1) in the first (or second) equation, then it is said that x_2(or x_1) Granger causes x_1(or x_2). The usefulness of the Granger causality lies in the fact that it is easy to use and hence since its development has been used in many different applications (7).

On the other hand, Shannon had defined the entropy of a random variable and developed the field of information theory. However, the information theory, as developed by Shannon, was symmetric and hence one cannot use it directly to infer causality. The concept of mutual information between two random variable x and y gives the *information* shared by x and y, but it does not give the direction of flow of information. In order to incorporate the directional sense in information theory, Marko developed what is now known as bi-directional information. It is a natural extension of the concept of mutual information between two random variables and furthermore, it also incorporates the sense of direction. Thus came the idea that elements of information theory can be used to infer the direction of information flow and hence causality. The major breakthrough, however, was achieved by Massey

(11) and later by Kramer (12) who generalized the concept of bi-directional information and defined what is known as directed information.

Let $X^n = \{X_1, X_2, \cdots, X_n\}$ and $Y^n = \{Y_1, Y_2, \cdots, Y_n\}$ be two stochastic processes, viewed as a sequence of random variables. Then the entropy (H) of the random variable X is defined as

$$H(X^n) = - \int_{\Omega_X} \rho(X^n) \log \rho(X^n) dX^n \tag{1.1}$$

where Ω_X is the event space of X and $\rho(X^n)$ is the probability distribution of X^n. Massey and Kramer (12) defined the directed information from X^n to Y^n as

$$I(X^n \to Y^n) = H(Y^n) - H(Y^n \parallel X^n) \tag{1.2}$$

where $H(Y^n)$ is the entropy of the sequence Y^n and

$$H(Y^n \parallel X^n) := \sum_{i=1}^{n} H(Y_i|Y^{i-1}, X^i) \tag{1.3}$$

is the entropy of Y^n *causally* conditioned on X^n. The idea of *causal conditioning* takes into account the direction of information flow and hence throws light on the causal structure of the stochastic process.

The concept of directed information is more suited to stochastic processes and in 2000, Thomas Schreiber proposed the idea of transfer entropy (9) which is geared to study the information flow in systems evolving in time. Transfer entropy is a difference of two conditional entropies and for Markov chains of order one, it is equal to conditional mutual information. The concept of transfer entropy is discussed in chapter 2.

However, though transfer entropy was geared more from system theoretic point of view, it fails to capture the zero influence in some cases (13). Liang and Kleeman (10) came up with a new formulation of information transfer in dynamical systems and showed that it captures the notion of zero transfer. To define the Liang-Kleeman information flow from state x to state y, they considered two different systems, namely, the original system and a modified system, where they froze x and propagated the modified system by one-time step to determine the amount of information flow from x to y. The Liang-Kleeman information flow is also explained in detail in chapter 2.

In (10), the authors had defined Liang-Kleeman information transfer in a two state system. In (14), their definition of information transfer was generalised to define information transfer between

subspaces of dynamical systems with additive noise. Liang and Kleeman carried their research forward and studied the information transfer in general n-dimensional continuous time and discrete time dynamical systems (15; 16; 17). They also studied the information transfer in two dimensional discrete time stochastic systems with additive noise (18). In (10), they also show that their formalism of information transfer is qualitatively consistent with the notions of transfer entropy and Granger causality. So Liang-Kleeman information transfer is also used to study the causal structure in a dynamical system. But the Liang-Kleeman information transfer cannot account for indirect transfer (17). Again, transfer entropy, though gives a measure of identifying the direction of information flow, and hence influence, it also suffers from some drawbacks. For example, it may not identify indirect influence and also may give erroneous results in presence of dominant neighbours (13). Directed information, as formulated by Massey and Kramer also suffers from similar deficiencies.

On the other hand, various researchers have tried to connect information theory and control theory. In (19), the authors have studied the classical concepts of controllability and observability from the information theoretic view point. Again, in (20), the authors have shown that the mutual information between the state and the estimate is a maximum in a discrete time Kalman filter. Moreover, in (21), the authors have studied the information flow in a Kalman-Bucy filter.

The main motivation for our work was to come up with a measure of causality which can identify the indirect links and also should be able to correctly infer about the zero influence. Moreover, it would also be nice if the measure of information transfer, so defined to capture the influence structure between the states in a dynamical system, can be extended to define information transfer between the various signals involved in a control system. In that case, it may be possible to connect the information transfer measure with some classic control theoretic notions. In that case, one may be able to connect the two fields of information theory and control theory and study them on a common base.

1.2 Our Contribution

Formalism for information transfer developed in this work is in dynamical systems setting and is closely related to and inspired from information transfer framework developed in (14; 10; 17) for nonlinear dynamical systems. We used the ideas of Liang-Kleeman transfer to propose an axiomatic

definition of information transfer in discrete linear dynamic network (22). The axioms were physically motivated and it was shown in (22) that there exists a unique expression for information transfer in a dynamic network satisfying these axioms. In this work, we use the ideas of directed information and transfer entropy and the intuitions of (10; 22) to define the information transfer.

In particular, instead of absolute entropy, we use the conditional entropy to define the information transfer. This definition is a natural extension of directed information and transfer entropy and we show that the new definition of information transfer satisfies the intuitions of information transfer, namely a) Zero transfer, b) Transfer asymmetry and c) Information conservation. The advantage of using conditional entropy to define information transfer lies in the fact that this new definition of one step information transfer can be naturally extended to define n-step transfer and average information transfer, which can be related to entropy rate of a variable and we find that the extension of the one step information transfer definition to n-step information transfer definition is the precise concept needed to capture indirect influence. This is one of the important consequence of our definition of information transfer. The n-step information transfer is not only a natural generalization of the one step information transfer, but the n-step information transfer serves as a natural measure of causality and we show, via examples, that in dynamic networks, where directed information (and transfer entropy) may give erroneous counter-intuitive results, as far as the causality structure is concerned, our definition of causality does in fact capture the true causal structure of the network.

The information transfer defined here-in is not a stand-alone concept and can be related to already existing concepts in control theory. In particular, a generalized definition of information transfer is provided to define information transfer between the various signals involved in a control dynamical system. With this generalized definition, we show that in a linear feedback control system, our information transfer is related to the Bode integral of the sensitivity transfer function from the output to the input. This result also establishes a connection between the average directed information and our information transfer.

Moreover, directed information and transfer entropy do not have any relation with the very basic and important control theoretic concepts like controllability and observability. In fact, they are also not related to the weaker notions of structural controllability and structural observability. This is

another serious drawback of using directed information and transfer entropy in a control theoretic setting. However, we show that our definition of information transfer can be connected to structural controllability and structural observability. In particular, we show that if the information transfer from the input to all the states is non-zero over any finite time step, then the system is structurally controllable. The result for structural observability is similar, where one has to look at the information transfer from the states to the output.

We also use the information transfer to define influence in a dynamic network. This is one of the more important applications of the proposed framework because defining influence and finding the influential nodes in a dynamic network has been and still is a problem of utmost importance in social media analysis, gene regulatory networks, neuroscience etc. We study two different real-life networks, where we use the proposed definition of information transfer to characterize influence. Further, we characterize stability in a dynamical system and identify the state (subspace) which are responsible for instability of the system. This has applications in power networks, where one can use the information transfer measure to classify voltage and angle instability. We demonstrate this on three different power networks, namely IEEE 3 bus system, IEEE 9 bus system and IEEE 39 bus system. The stability analysis of power systems from information transfer viewpoint served as a motivation for studying phase transitions in complex systems and we show that information transfer can predict phase transition in complex systems as well.

We also provide algorithms to compute information transfer from time series data. We provide a general algorithm for general non-linear maps and also provide a different algorithm, based on ideas from robust optimization, for information transfer computation from data coming from linear systems. We show how the algorithms can be used for topology identification of linear networks and also, as another application we analyze stock market and characterize influence in the same, using stock market data.

1.3 Organization of the Report

The report is organized as follows. In chapter 2, we develop the basic concepts and definitions needed to define the information transfer in a discrete time dynamical system. In chapter 3, we define

the notion of information transfer and prove the properties that it satisfies. We also define the n-step transfer in this section and analyze some of its properties. We further provide a set-oriented method for finding the information transfer. Chapter 4 is dedicated to finding the information transfer in linear dynamical control systems. Chapter 5 is dedicated to discussing how our definition of information transfer gives the correct causal structure in a dynamical system. In chapter 6, we discuss some basic examples where we use the information transfer to characterize influence. In chapter 7, we use information transfer to characterize influence and stability in power systems and in chapter 8 we use these intuitions to analyze phase transitions using information transfer measure. In chapter 9, we discuss two algorithms for computing information transfer from time series data and in chapter 10, we analyze the causal structure of US stock market using stock price data. Finally, we conclude the report in chapter 11.

CHAPTER 2. PRELIMINARIES FOR INFORMATION TRANSFER FORMULATION

In this chapter, we review the basics of information theory and the existing notions of information flow. Throughout this work, by information of a random variable, we will mean the Shannon entropy of the random variable. So, if X is a random variable, with probability distribution $\rho(x)$, defined over the event space Ω, then the entropy of X is given by

$$H(X) = -\int_\Omega \rho(x) \log \rho(x) dx \tag{2.1}$$

This notion of entropy of a random variable was defined by Shannon (23) and physically it captures the uncertainty of the random variable X. This is also known as the information contained in the random variable X. In our formulation of information transfer, we study how this *information* is flowing between the states of a dynamical system. So the fundamental problem is to study how the probability density functions evolve under the system dynamics. We consider the following dynamical system (2.2) and address the problem of information transfer between the states of the system.

$$z(t+1) = f(z(t)) + \xi(t) \tag{2.2}$$

where $z(t) \in \mathbb{R}^N$, $\xi(t) \in \mathbb{R}^N$ is assumed to vector-valued random variable and $\xi(0), \xi(1), \ldots$ are independent random vectors each having the same density g. The mapping $f : \mathbb{R}^N \to \mathbb{R}^N$ is assumed to be at least continuous. Let $z = (z_1, \ldots, z_N)^\top \in \mathbb{R}^N$. We are interested in finding the information transfer from state z_i to state z_j, as the system evolves from time step t to time step $t+1$. We denote this transfer by the notation $[T_{z_i \to z_j}]_t^{t+1}$. More generally, we are also interested in deriving the information transfer between the subspaces of the dynamical system.

Let $\rho(z, t)$ be probability density function at time t. The propagation of probability density function under system dynamics (2.2) is described by the linear Perron-Frobenius operator $\mathbb{P} : \mathcal{L}^1 \to \mathcal{L}^1$

(24).

2.1   Perron-Frobenius Operator

Consider a discrete-time process defined by

$$x(n+1) = f(x(n)) + \xi(n)$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a measurable, not necessarily nonsingular, transformation and $\xi(0), \xi(1), \cdots$ are independent random vectors each having the same density g. Let the density of $x(n)$ be denoted by ρ_n. Let $h : \mathbb{R}^d \to \mathbb{R}$ be an arbitrary, bounded, measurable function. The expectation of $h(x(n+1))$ is

$$E(h(x(n+1))) = \int_{\mathbb{R}^d} h(x)\rho_{n+1}(x)dx$$

Because of the fact that the joint density of $(x(n), \xi(n))$ is just $\rho_n g$, we have

$$
\begin{aligned}
E(h(x(n+1))) &= E[h\left(f(x(n) + \xi(n))\right)] \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(f(y) + z)\rho_n(y)g(z)dydz
\end{aligned}
$$

By a change of variables,

$$E(h(x(n+1))) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h(x)\rho_n(y)g(x - f(y))dxdy$$

Using the fact that h was an arbitrary, bounded, measurable function, we have

$$\rho_{n+1} = \int_{\mathbb{R}^d} \rho_n(y)g(x - f(y))dy \tag{2.3}$$

From the above equation, we define an operator $\mathbb{P} : L^1 \to L^1$ by

$$\mathbb{P}\rho(x) = \int_{\mathbb{R}^d} \rho(y)g(x - f(y))dy \tag{2.4}$$

This operator $\mathbb{P}$ is a Markov operator and dictates the evolution of the density and is the Perron-Frobenius operator for the system defined by (2.2). Hence, under the system dynamics (2.2), the evolution of the density function $\rho(z(t))$ is given by

$$\rho_{t+1}(z) = [\mathbb{P}\rho_t](z) = \int_{\mathbb{R}^N} \rho_t(\tilde{z})g(z - f(\tilde{z}))d\tilde{z} \tag{2.5}$$

Remark 1. *For convenience of notation, we will denote the density $\rho_t(z)$ by $\rho(z(t))$. Also at places we will use $'$ notation to denote $t+1$ time instant of state variable i.e., $z' := z(t+1)$ and $z(t) := z$.*

2.2 Directed Information, Transfer Entropy and Liang-Kleeman Information Transfer

Shannon's information theory is symmetric and hence does not capture the causal structure in multivariate time series or in a dynamical system. To incorporate the sense of direction, Marko proposed, what is known today as bi-directional information (25). Massey and Kramer (11; 12) generalized the concept of bi-directional information and defined directed information between two sequences of random variables and thus incorporated the sense of direction in Shannon's information theory.

2.2.1 Directed Information

Let $X^n = \{X_1, X_2, \cdots, X_n\}$ and $Y^n = \{Y_1, Y_2, \cdots, Y_n\}$ be two stochastic processes, viewed as a sequence of random variables. Then the entropy (H) of the random variable X is defined as

$$H(X^n) = -\int_{\Omega_X} \rho(X^n) \log \rho(X^n) dX^n \tag{2.6}$$

where Ω_X is the event space of X and $\rho(X^n)$ is the probability distribution of X^n. Given the distributions $\rho(X^n)$ and $\rho(Y^n)$, the mutual information between X and Y is defined as

$$\begin{aligned}
I(X^n; Y^n) &= \int_{\Omega_X \times \Omega_Y} \rho(X^n Y^n) \log \frac{\rho(X^n Y^n)}{\rho(X^n)\rho(Y^n)} \\
&= H(Y^n) - H(Y^n|X^n)
\end{aligned} \tag{2.7}$$

where $H(Y^n|X^n)$ is the conditional entropy of Y^n, conditioned on X^n, such that

$$\rho(Y^n|X^n) = \prod_{i=1}^{n} \rho(Y_i|Y^{i-1}, X^n) \tag{2.8}$$

Note that each term of the product on the right hand side of equation (2.8) depends on the entire entire sequence of X. Massey and Kramer (12) defined the directed information from X^n to Y^n as

$$I(X^n \to Y^n) = H(Y^n) - H(Y^n \parallel X^n) \tag{2.9}$$

where $H(Y^n)$ is the entropy of the sequence Y^n and

$$H(Y^n \parallel X^n) := \sum_{i=1}^{n} H(Y_i|Y^{i-1}, X^i) \tag{2.10}$$

is the entropy of Y^n *causally* conditioned on X^n. This differs from

$$H(y^N|x^N) = \sum_{n=1}^{N} H(y_n|y^{n-1}x^N)$$

in the fact that x^n is replaced by x^N.

Directed information is not symmetric in X and Y and thus depicts the direction of information flow between the stochastic processes. This observation led to the use of directed information to infer about the causal structure in a multivariate stochastic process.

2.2.2 Transfer Entropy

In system theoretic perspective, Thomas Schreiber (26) came up with an information theoretic measure, called *transfer entropy*, to distinguish driving and responding elements in a dynamical system and also to detect asymmetry in the interaction of the subsystems. The transfer entropy is defined as follows.

Definition 2 (Transfer Entropy). *: Let $\rho(\cdot)$ denote the probability density of the appropriate random variable, then in the case of Markov chain of order one, the transfer entropy from z_i to z_j, at time step t, is*

$$T^S_{z_i \to z_j} = \sum \rho(z_j^{t+1}, z_j^t, z_i^t) \log \frac{\rho(z_j^{t+1}|z_j^t, z_i^t)}{\rho(z_j^{t+1}|z_j^t)} \tag{2.11}$$

The transfer entropy is a difference of two conditional entropies and its non-symmetric nature captures the direction of information flow in a dynamical system. we note that for Markov chain of order one, transfer entropy is the conditional mutual information between z_j^{t+1} and z_j^t, given z_i^t. For the special case of vector auto-regressive processes, transfer entropy can be shown to be equivalent to Granger causality (27) and hence, when Granger causality analysis is difficult to carry out, often transfer entropy is used to infer about the causal structure.

2.2.3 Liang-Kleeman Information Transfer

Both directed information and transfer entropy were motivated from the mutual information and in both the cases the definition of mutual information was modified to incorporate a sense of direction. Liang and Kleeman (10) introduced a new way of defining information transfer in a dynamical system. The intuition they had is that the information transfer from a state z_i to z_j is the difference of the entropy of z_j when z_i is present in the dynamics and when z_i is absent from the dynamics. So, in order to find the Liang-Kleeman information flow from z_i to z_j, as the system (2.2) evolves from time step t to time step $t + 1$, one has to compute the entropy of two different systems, one of them being the original system and the other one being a modified system, where z_i is held *frozen* from time step t to time step $t + 1$. We briefly describe this procedure for a two dimensional system.

Let

$$z_1(t + 1) = f_1(z_1(t), z_2(t))$$
$$z_2(t + 1) = f_2(z_1(t), z_2(t))$$
(2.12)

be a two dimensional discrete time system. Then the Liang-Kleeman information transfer from the state z_1 to z_2, as the system evolves from time step t to time step $t + 1$ is

$$T_{z_1 \to z_2}^{LK} = H(z_2(t + 1)) - H_{\not z_1}(z_2(t + 1))$$
(2.13)

where $H_{\not z_1}(z_2(t + 1))$ is the entropy of z_2 at time step $t + 1$, where z_1 has been held frozen. That is, to calculate $H_{\not z_1}(z_2(t + 1))$, one has to consider the reduced system

$$z_2(t + 1) = f_2(z_1(t), z_2(t)) + \xi_2(t)$$
(2.14)

where $z_1(t)$ is treated as a parameter.

Liang-Kleeman information transfer definition is intuitive and captures the important notion of zero transfer, that is, when z_i does not appear in the dynamics of z_j, then the information transfer from z_i to z_j is zero. However, one serious drawback of Liang-Klemman information transfer is that, like directed information and transfer entropy, it can not capture the indirect transfer. By this, we mean that it may happen that in a dynamical system, z_i does not affect the dynamics of z_k directly, but does so via z_j. That is, z_i affects the dynamics of z_j which in turn affects z_k. So indirectly, z_i affects

z_k and so there should be a non-zero information flow from z_i to z_k. This indirect transfer is not captured by the Liang-Kleeman information flow. Directed information and transfer entropy, in this case, will give a non-zero transfer, but it fails to identify whether the transfer was direct or via some other state. Moreover, there are some other issues with directed information and transfer entropy, as far as the causal structure is concerned. We will discuss them in details in chapter 5.

With this, now we are in a position to introduce our notion of information transfer.

CHAPTER 3. INFORMATION TRANSFER

In this chapter, we provide the definition of information transfer in a dynamical system and discuss its basic properties. We first define the one step information transfer and discuss it in details and then we will generalize the definition of one step information transfer to define information transfer over n-steps.

3.1   One Step Information Transfer

To derive the expression for information transfer from state z_i to z_j of the system given by (2.2), we first split the state $z \in \mathbb{R}^N$ into two subspace. In particular, $z = (x^\top, y^\top)^\top$, where $x \in \mathbb{R}^{|x|}$ and $y \in \mathbb{R}^{|y|}$ with $|x|$ denotes the cardinality of space x and hence $|x| + |y| = N$. We first present an expression for information transfer for the two subspace case i.e., $T_{x \to y}$. Towards this goal we write the system dynamics in terms of x and y subspace as follows:

$$
\begin{aligned}
z(t+1) &= \begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} f_x(x,y) \\ f_y(x,y) \end{pmatrix} + \begin{pmatrix} \xi_x(t) \\ \xi_y(t) \end{pmatrix} \\
&:= f(x(t), y(t)) + \xi(t)
\end{aligned} \tag{3.1}
$$

We later show how the expression for information transfer between two subspace generalizes to systems with more than two subspace. There are existing concepts, like directed information (11), transfer entropy (26) and Liang-Kleeman information (10), which study the flow of information in a dynamical system. However, these concepts do not capture all the intuitions for information transfer. For example, while all the above three concepts satisfy the information transfer asymmetry property, only Liang-Kleeman information transfer satisfies the zero transfer property. However, the Liang-Kleeman information does not satisfy the information conservation property. We show that our

proposed definition of information transfer satisfies all the three properties of information transfer, namely,

1. Zero information transfer.

2. Information transfer asymmetry.

3. Information conservation.

The property of information conservation is unique to our formulation of information transfer in dynamical system. We will discuss the above three properties of information transfer in details later.

In order to study the information transfer between the subspaces of the dynamical system (2.2), it is sufficient to study the transfer from subspace x to subspace y, for the system (3.1). To calculate the information transfer from x to y, we need to study the evolution of the density, not only for (3.1), but also for the modified system, given by

$$
\begin{aligned}
x^{'} &= x \\
y^{'} &= f_y(x, y) + \xi_y
\end{aligned}
\tag{3.2}
$$

Following (10), we say that x is frozen (held constant), as the system evolves from time step t to time step $t + 1$.

Remark 3. *We use the notation $\rho(\cdot)$ to denote the density for any variable for the system (3.1) and $\rho_{\not{x}}(\cdot)$ to denote the density of a variable for the system (3.2). Here $\not{x}$ denotes that x has been frozen.*

With this, we will now present some definitions which will be important for understanding the properties of information transfer.

Definition 4. *y dynamics is said to be not influenced by dynamics of x if*

$$
\rho(y(t+1)|y(t)) = \rho_{\not{x}}(y(t+1)|y(t))
$$

The above condition implies that the conditional probability density of $y(t + 1)$, conditioned on $y(t)$, is not changed by the absence of $x(t)$ and hence the dynamics of y are not influenced by x.

Definition 5. *Consider the following splitting of $y \in \mathbb{R}^{|y|}$ subspace i.e., $y = (y_1^\top, y_2^\top)^\top$, where $y_k \in \mathbb{R}^{|y_k|}$ for $k = 1, 2$ and $|y_1| + |y_2| = |y|$. We say that y_1 and y_2 are dynamically independent if*

$$\rho(y) = \rho(y_1)\rho(y_2)$$

$$\Downarrow$$

$$\rho(y'|y) = \rho(y_1'|y_1)\rho(y_2'|y_2) \text{ and } \rho_{\not x}(y'|y) = \rho_{\not x}(y_1'|y_1)\rho_{\not x}(y_2'|y_2)$$

where $y' = y(t+1), y = y(t)$, and $x = x(t)$.

The above definitions of zero influence and dynamical independence are used in stating the following three properties for information transfer based on which we provide the expression for information transfer :

Properties :

1. Zero transfer: If the y dynamics is not influenced by x dynamics then the information transfer from $x \rightarrow y$ i.e., $T_{x \rightarrow y} = 0$.

2. Transfer Asymmetry: If the y dynamics is not influenced by x dynamics but x dynamics is influenced by y dynamics then $T_{x \rightarrow y} = 0$ but $T_{y \rightarrow x} \neq 0$.

3. Information conservation : If y_1 and y_2 are dynamically independent as per Definition 5, then following conservation property for the information transfer holds true

$$T_{x \rightarrow y}(t) = T_{x \rightarrow y_1}(t) + T_{x \rightarrow y_2}(t)$$

Our goal is to derive an expression for information transfer that satisfies the above properties of transfer. We provide following definition for information transfer for dynamical system

Definition 6. *[Information Transfer] The information transfer between x and y for dynamical system (2.2) as the system evolves from time t to time $t+1$ and denoted by $[T_{x \rightarrow y}]_t^{t+1}$ is given by following formula*

$$[T_{x \rightarrow y}]_t^{t+1} = H(\rho(y(t+1)|y(t))) - H(\rho_{\not x}(y(t+1)|y(t))) \tag{3.3}$$

where $H(\rho(y)) = -\int_{\mathbb{R}^{|y|}} \rho(y) \log \rho(y) dy$ *is the entropy of probability density function* $\rho(y)$ *and* $H(\rho_{\cancel{x}}(y(t+1)|y(t))$ *is the entropy of* $y(t+1)$, *conditioned on* $y(t)$, *where* x *has been freezed. The two entropies in above expression are defined as follows:*

$$H(\rho(y(t+1)|y(t))) = \int_{\mathbb{R}^{2|y|}} \rho(y'|y) \log \rho(y'|y) dy' dy$$

$$H(\rho_{\cancel{x}}(y(t+1)|y(t))) = \int_{\mathbb{R}^{2|y| \times |x|}} \rho_{\cancel{x}}(y'|y) \log \rho_{\cancel{x}}(y'|y) \rho(x) dx dy' dy$$

The above definition of information transfer can be understood by rewriting the expression of information transfer as follows:

$$H(\rho(y(t+1)|y(t))) = [T_{x \to y}]_t^{t+1} + H(\rho_{\cancel{x}}(y(t+1)|y(t))) \tag{3.4}$$

so that the total entropy of y is the sum of transfer from x and the entropy of y, when x is absent. Now, using

$$H(Y|X) = H(X,Y) - H(Y)$$

and the fact that since x is frozen at time t, we have that $H(y(t)) = H_{\cancel{x}}(y(t))$. Hence, we have

$$[T_{x \to y}]_t^{t+1} = H(\rho(y(t+1), y(t))) - H_{\cancel{x}}(\rho(y(t+1), y(t))) \tag{3.5}$$

In the following, we prove that the proposed definition of information transfer satisfies the properties of information transfer

Theorem 7. *The proposed expression of information transfer Eq. (3.3) satisfies all the three properties of information transfer.*

Proof.   1. Zero transfer: Assume that dynamics of y is not influenced by x, we then have $\rho(y(t+1)|y(t)) = \rho_{\cancel{x}}(y(t+1)|y(t))$. Hence, we have

$$T_{x \to y}(t) = H(\rho(y(t+1)|y(t))) - H(\rho(y(t+1)|y(t))) = 0 \tag{3.6}$$

2. Transfer asymmetry: Assume that y is not influenced by x but y is influencing x, we then have

$$\rho(y(t+1)|y(t)) = \rho_{\cancel{x}}(y(t+1)|y(t))$$

$$\rho(x(t+1)|x(t)) \neq \rho_{\cancel{y}}(x(t+1)|x(t)).$$

From the formula for information transfer, we get

$$T_{x \to y}(t) = 0, \quad T_{y \to x}(t) \neq 0.$$

3. Information conservation: Let y_1 and y_2 be dynamically independent. Then we have

$$\rho(y) = \rho(y_1)\rho(y_2)$$

$$\rho(y'|y) = \rho(y_1'|y_1)\rho(y_2'|y_2)$$

and $\qquad \rho_{\cancel{x}}(y'|y) = \rho_{\cancel{x}}(y_1'|y_1)\rho_{\cancel{x}}(y_2'|y_2)$

Hence, $H(\rho(y'|y)) = H(\rho(y_1'|y_1)) + H(\rho(y_2'|y_2))$ and $H(\rho_{\cancel{x}}(y'|y)) = H(\rho_{\cancel{x}}(y_1'|y_1)) + H(\rho_{\cancel{x}}(y_2'|y_2))$,

so that

$$
\begin{aligned}
[T_{x \to y}]_t^{t+1} &= H(\rho(y'|y)) - H(\rho_{\cancel{x}}(y'|y)) \\
&= [H(\rho(y_1'|y_1)) - H(\rho_{\cancel{x}}(y_1'|y_1))] + [H(\rho(y_2'|y_2)) - H(\rho_{\cancel{x}}(y_2'|y_2))] \\
&= [T_{x \to y_1}]_t^{t+1} + [T_{x \to y_2}]_t^{t+1}
\end{aligned}
$$

$\square$

Equation (3.3) gives the information transfer from subspace x to subspace y of the dynamical system (3.1), as it evolves from time t to time $t+1$. The general expression for information transfer between scalar state z_i to scalar state z_j can be written as

$$[T_{z_i \to z_j}]_t^{t+1} = H(\rho(z_j(t+1)|z_j(t))) - H(\rho_{\cancel{z_i}}(z_j(t+1)|z_j(t))) \tag{3.7}$$

Example 8. Correlation does not imply causation. *It is well known that correlation may not imply causation. Our definition of information transfer also verifies this fact, as is illustrated by the following example. Consider the following linear system with additive zero mean unit variance i.i.d. Gaussian noise.*

$$
\begin{aligned}
x(t+1) &= .5x(t) + \xi_x(t) \\
y(t+1) &= .3x(t) + .7y(t) + \xi_y(t)
\end{aligned}
\tag{3.8}
$$
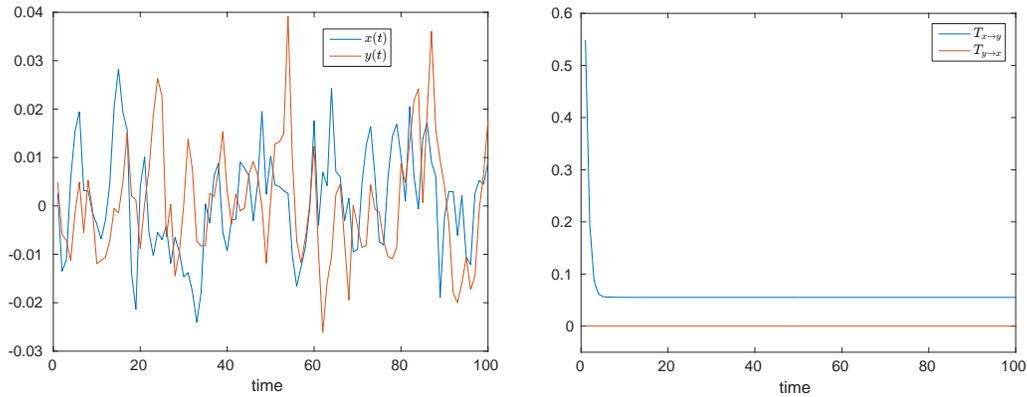
Figure 3.1   (a) The states are highly correlated. (b) The information transfer between the states.

As can be seen from fig 8(a), the states x and y are highly correlated, but from the system dynamics (3.8), the state y does not affect the dynamics of x and so, intuitively, y should not cause x. This is reflected in fig. 8(b), where we find that the information transfer from y to x is zero, whereas, since x affects y, the transfer from x to y is non-zero.

The above example allows us to believe that our formalism of information transfer can be used to infer about the causal structure in a dynamical system. In later chapters, we will see, when directed information and transfer entropy gives counter-intuitive results about the causal structure in a dynamical system, how our formalism gives the correct information about the causal structure.

3.1.1   Information Transfer and Causal Inference in Nonlinear Systems

In this subsection, we connect information transfer and causal inference for general nonlinear systems. However, we will consider systems without noise. In particular, we consider the system

$$z_{t+1} = S(z_t) \tag{3.9}$$

where $S : \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be at least continuous and $z(t) \in \mathbb{R}^n$. In component form, equation (3.9) can be written as

$$
\begin{aligned}
z_{t+1}^1 &= S_1(z_t^1, z_t^2, \cdots, z_t^n) \\
z_{t+1}^2 &= S_2(z_t^1, z_t^2, \cdots, z_t^n) \\
&\vdots \\
z_{t+1}^n &= S_n(z_t^1, z_t^2, \cdots, z_t^n)
\end{aligned}
\tag{3.10}
$$

Information transfer between the states of a dynamical system is defined in terms of difference of certain entropies, as the dynamical system evolves in time. As such to compute the information transfer, it is necessary to know the evolution of densities. This is given by the Perron-Frobenius (PF) operator

$$
\mathbb{P} : \mathcal{L}^1(\mathbb{R}^n) \to \mathcal{L}^1(\mathbb{R}^n)
\tag{3.11}
$$

and is defined as

$$
\int_\omega \mathbb{P}\rho(z)dz = \int_{S^{-1}(\omega)} \rho(z)dz
\tag{3.12}
$$

for any $\omega \in \mathcal{B}(\mathbb{R}^n)$, where $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ-algebra on $\mathbb{R}^n$. When S is non-singular and invertible, the PF operator can be written explicitly as

$$
\mathbb{P}\rho(z) = \rho[S^{-1}(z)]|J^{-1}|
\tag{3.13}
$$

where J is the Jacobian of S and $|\cdot|$ signify the determinant. The joint density $\rho(z)$ is used to define the Shannon entropy

$$
H(z) = -\int_\Omega \rho(z) \log \rho(z)dz
\tag{3.14}
$$

where Ω is the sample space.

Assumption 9. *We assume Ω to be Cartesian product of $\Omega_1, \Omega_2, \cdots, \Omega_n$, in which $z^1, z^2, \cdots, z^n$ are respectively lying. Here each Ω_i are open in $\mathbb{R}$, for $i = 1, 2, \cdots, n$.*

The marginal entropy of any variable of interest, say z_1 is defined as

$$
H(z^1) = -\int_{\Omega_1} \rho_1(z^1) \log \rho_1(z^1)dz^1
\tag{3.15}
$$

where $\rho_1 = \rho(z^1) = \int_{\Omega_{2n}} \rho(z) dz^2 dz^3 \cdots dz^n$ is the marginal distribution of z^1. Here we use the notation

$$\Omega_{jn} = \Omega_j \times \Omega_{j+1} \times \cdots \times \Omega_n, \quad 1 \leq j \leq n, \text{ and } \Omega_{1n} = \Omega. \tag{3.16}$$

We further use the notation

$$\rho_{i\!\!/} = \rho(z^1, \cdots z^{i-1}, z^{i+1}, \cdots, z^n)$$
$$= \int_{\Omega_i} \rho(z^1, \cdots z^{i-1}, z^{i+1}, \cdots, z^n) dz^i$$
$$\rho_{i\!\!/j\!\!/} = \rho(z^1, \cdots z^{i-1}, z^{i+1}, \cdots, z^{j-1}, z^{j+1} \cdots, z^n)$$
$$= \int_{\Omega_i \times \Omega_j} \rho(z^1, \cdots z^{i-1}, z^{i+1}, \cdots, z^{j-1}, z^{j+1}, \cdots, z^n) dz^i dz^j$$

With this, we state the following theorem which connects information transfer with causal inference.

Theorem 10. *If S_j is independent of z^i, then $T_{z^i \to z^j} = 0$. At the same time, $T_{z^i \to z^j}$ need not be zero if S_j does not rely on z^i.*

Proof. Without loss of generality, we assume $i = 2$ and $j = 1$, so that we look at the information transfer from z^2 to z^1. We consider the system $z_{t+1} = S(z_t)$ given in component form as in (3.10). Define $\hat{z}_t^i = z_{t+1}^i$ and consider the dynamical system

$$z_{t+1}^i = S_i(z)$$
$$\hat{z}_{t+1}^i = S_i(\hat{z}) \tag{3.17}$$

for $i = 1, 2, \cdots, n$. Now, from (3.5),

$$\begin{aligned} T_{z^2 \to z^1} &= H(z_{t+1}^1, z_t^1) - H_{\!\!/2}(z_{t+1}^1, z_t^1) \\ &= H(z_{t+1}^1, \hat{z}_{t+1}^1) - H_{\!\!/2}(z_{t+1}^1, \hat{z}_{t+1}^1) \end{aligned} \tag{3.18}$$

It is given that $S_1(z)$ is independent of z^2. Hence, $S_1(\hat{z})$ is independent of $\hat{z}^2$ and also z^2. To prove that if $S_1(z)$ is independent of z_2, then $T_{z_2 \to z_1} = 0$, we have to show

$$H(z_{t+1}^1, \hat{z}_{t+1}^1) = H_{\!\!/2}(z_{t+1}^1, \hat{z}_{t+1}^1)$$

Let U be the sample space where the system (3.17) evolves. Note that U satisfies assumption (9) and we use the notation defined in (3.16). For notational convenience we use the notation x_t^1 to signify the subspace $(z_t^1, \hat{z}_t^1)^\top$, $x_t^2 = z_t^2$, $y_t = \hat{z}_t^2$, $x_t^i = (z_t^i, \hat{z}_t^i)^\top$ for $i = 3, \cdots, n$. Similarly we define U_i and U_y. Hence we need to prove

$$H(x_{t+1}^1) = H_{\not{2}}(x_{t+1}^1)$$

The dynamical system (3.17) can be written as

$$
\begin{aligned}
x_{t+1}^i &= \phi_i(x_t, y_t) \\
x_{t+1}^2 &= \phi_2(x_t, y_t) \\
y_{t+1} &= \phi_y(x_t, y_t)
\end{aligned}
\tag{3.19}
$$

for $i = 1, 3, \cdots, n$ and $\phi_i = (S_i, S_i)^\top$, $\phi_2 = S_2$ and $\phi_y = S_2$. Define $\Phi = \begin{pmatrix} \phi_i^\top & \phi_2 & \phi_y \end{pmatrix}^\top$. For any subset of $\tilde{U}_1$, $\tilde{u}_1 \in \tilde{U}_1$, we have

$$
\begin{aligned}
\int_{\tilde{u}_1} (\mathbb{P}\rho_{\not{2}})_1(x^1) dx^1 &= \int_{\tilde{u}_1 \times U_{3n} \times U_y} (\mathbb{P}\rho_{\not{2}})(x^1, y, x^3, \cdots, x^n) \\
&\quad \times dx^1 dx^3 \cdots dx^n dy \\
&= \int_{\Phi_{\not{2}}^{-1}(\tilde{u}_1 \times U_{3n} \times U_y)} \rho_{\not{2}}(x^1, y, x^3, \cdots, x^n) \\
&\quad \times dx^1 dx^3 \cdots dx^n dy
\end{aligned}
$$

Now[1],

$$\Phi_{\not{2}}^{-1}(\tilde{u}_1 \times U_{3n} \times U_y) = \Phi_{1\not{2}}^{-1}\tilde{u}_1 \times U_{3n} \times U_y \tag{3.20}$$

Hence,

$$
\begin{aligned}
\int_{\tilde{u}_1} (\mathbb{P}\rho_{\not{2}})_1(x^1) dx^1 &= \int_{\Phi_{1\not{2}}^{-1}\tilde{u}_1} dx^1 \cdot \int_{U_{3n} \times U_y} \rho_{\not{2}}(x^1, x^3, \cdots, x^n, y) \\
&\quad \times dx^1 dx^3 \cdots dx^n dy \\
&= \int_{\Phi_{1\not{2}}^{-1}\tilde{u}_1} \rho_1(x^1) dx^1 = \int_{\phi_1^{-1}\tilde{u}_1} \rho_1(x^1) dx^1
\end{aligned}
$$

---

[1]In the proof, $\Phi_{1\not{2}}^{-1}\tilde{u}_1$ and $\phi_1^{-1}\tilde{u}_1$ is the projection of the Cartesian product on the subspace x^1.

where the last equality follows from the fact that ϕ_1 and hence ϕ_1^{-1} is independent of x_2.

Again,

$$
\begin{aligned}
\int_{\tilde{u}_1} (\mathbb{P}\rho)_1(x^1)dx^1 &= \int_{\tilde{u}_1 \times U_{2n} \times U_y} \mathbb{P}\rho(x,y)dxdy \\
&= \int_{\Phi^{-1}(\tilde{u}_1 \times U_{2n} \times U_y)} \rho(x,y)dxdy \\
&= \int_{\phi^{-1}\tilde{u}_1 \times U_{2n} \times U_y} \rho(x,y)dxdy \\
&= \int_{\phi_1^{-1}\tilde{u}_1} dx^1 \int_{U_{2n} \times U_y} \rho(x)dx^2 \cdots dx^n dy \\
&= \int_{\phi_1^{-1}\tilde{u}_1} \rho_1(x^1)dx^1.
\end{aligned}
$$

Hence,

$$
\int_{\tilde{u}_1} (\mathbb{P}\rho_{\not{x}^2})_1(x^1)dx^1 = \int_{\tilde{u}_1} (\mathbb{P}\rho)_1(x^1)dx^1 \tag{3.21}
$$

for any $\tilde{u}_1 \in U_1$. Hence, $(\mathbb{P}\rho_{\not{x}^2})_1 = (\mathbb{P}\rho)_1$ almost everywhere, when ϕ_1 is independent of x^2. Now, in (15) it was shown that

$$
\begin{aligned}
H_{\not{x}^2}(x^1_{t+1}) &= -\int_U (\mathbb{P}\rho_{\not{x}^2})_1(y^1) \log((\mathbb{P}\rho_{\not{x}^2})_1(y^1)) \\
&\quad \cdot \rho(x^2|x^1, x^3 \cdots, x^n, y) \\
&\quad \cdot \rho_{3\cdots n,y}(x^3, \cdots x^n, y)dy^1 dx^2 dx^3 \cdots dx^n dy \\
&= -\int_U (\mathbb{P}\rho)_1(y^1) \log((\mathbb{P}\rho)_1(y^1)) \\
&\quad \cdot \rho(x^2|x^1, x^3 \cdots, x^n, y) \\
&\quad \cdot \rho_{3\cdots n,y}(x^3, \cdots x^n, y)dy^1 dx^2 dx^3 \cdots dx^n dy \\
&= -\int_{U_1} (\mathbb{P}\rho)_1(y^1) \log((\mathbb{P}\rho)_1(y^1))dy^1 \\
&\quad \cdot \Big[\int_{U_2} \int_{U_{3n} \times U_y} \rho(x^2|x^1, x^3 \cdots, x^n, y) \\
&\quad \cdot \rho_{3\cdots n,y}(x^3, \cdots x^n, y)dx^2 dx^3 \cdots dx^n dy \Big] \\
&= -\int_{U_1} (\mathbb{P}\rho)_1(y^1) \log((\mathbb{P}\rho)_1(y^1))dy^1 \\
&= H(x^1_{t+1})
\end{aligned}
$$

This is because the integrand in the brackets integrate to one. Hence, $H(x^1_{t+1}) = H_{\not{x}^2}(x^1_{t+1})$, that is,

$$
H(z^1_{t+1}) = H_{\not{x}^2}(z^1_{t+1})
$$

Hence, $T_{z^2 \to z^1} = 0$. $\qquad \square$

3.2 n-step Information Transfer

The information transfer defined in definition (6) gives the information transferred from x to y as the dynamical system evolves from time step t to time step $t + 1$. So one can look at this definition as a *one*-step transfer. However, it may happen that x_1 does not affect x_2 in one time step, but affects it in two time steps. For, example, consider the linear system described by the graph given in figure 3.2.

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$
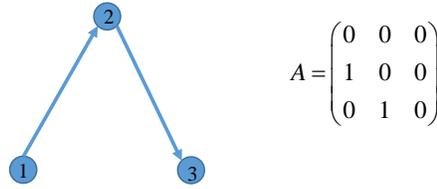
Figure 3.2   A system where two step transfer is non-zero

Here the one step information transfer from x_1 to x_3 is zero because x_1 does not affect the dynamics of x_3 in one time step. But x_3, at the second time step, is affected by x_1 and this should lead to a non-zero transfer. This is what we call indirect transfer.

We propose a natural generalization of the definition of one step information transfer to n step transfer. The n step transfer from x to y is defined as transfer from x to y as the system evolves from time step t to time step $t + n$ for $n > 0$. To define the n step transfer we introduce following notation

$$y_t^{t+n} = (y(t+n), y(t+n-1), \ldots, y(t)).$$

The n step transfer is defined as follows :

Definition 11. *[n-step Transfer] The information transferred over $n > 0$ time steps from x to y, denoted as $[T_{x \to y}]_t^{t+n}$, as the dynamical system (3.1) evolves from time step t to $t + n$ is defined as*

$$[T_{x \to y}]_t^{t+n} = H(\rho(y(t+n)|y_t^{t+n-1}) - H(\rho_{\mathscr{Y}_t^{t+n-1}}(y(t+n)|y_t^{t+n-1})$$

where $\rho_{\mathscr{Y}_t^{t+n-1}}(\cdot)$ is the notation that is used to signify the fact that dynamics in x component is freezed over the time period from t to $t + n - 1$.

Using the proposed definition of n-step transfer to the system of Fig. 3.2, we have the two step information transfer from x_1 to x_3 to be non zero. This is in accordance with the intuition that since x_1 affects x_3 in the second time step, the information transfer in two time step should be non-zero. With this definition of n-step transfer, we have

$$
\begin{aligned}
[T_{x \to y}]_t^{t+2} &= H(\rho(y(t+2)|y_t^{t+1}) - H(\rho_{\not x_t^{t+1}}(y(t+2)|y_t^{t+1})) \\
&= H(\rho(y_t^{t+2})) - H(\rho_{\not x_t^{t+1}}(y_t^{t+2})) - [T_{x \to y}]_t^{t+1}
\end{aligned}
$$

where we have used the fact that

$$
[T_{x \to y}]_t^{t+1} = H(\rho(y_t^{t+1})) - H(\rho_{\not x_t}(y_t^{t+1}))
$$

Hence,

$$
[T_{x \to y}]_t^{t+2} + [T_{x \to y}]_t^{t+1} = H(\rho(y_t^{t+2})) - H(\rho_{\not x_t^{t+1}}(y_t^{t+2})) \tag{3.22}
$$

Equation (3.22), says that the total change in entropy, due to x, of the joint distribution of $(y(t+2)y(t+1)y(t))$, as the system (3.1) evolves from time step t to time step $t+2$, is equal to the sum of the information transfers from x to y over one time step and two time steps. In general, we have the following

Theorem 12.

$$
\sum_{i=1}^{n} [T_{x \to y}]_t^{t+i} = H(\rho(y_t^{t+n})) - H(\rho_{\not x_t^{t+n-1}}(y_t^{t+n}))
$$

Proof. We prove this by induction. We saw that this is true for $i = 1, 2$. Let this be true for $i = k$. Now,

$$
\begin{aligned}
[T_{x \to y}]_t^{t+k+1} &= H(\rho(y(t+k+1)|y_t^{t+k})) \\
&\quad - H(\rho_{\not x_t^{t+k}}(y(t+k+1)|y_t^{t+k})) \\
&= \left(H(\rho(y_t^{t+k+1})) - H(\rho_{\not x_t^{t+k}}(y_t^{t+k+1})) \right) - \\
&\quad \left(H(\rho(y_t^{t+k})) - H(\rho_{\not x_t^{t+k-1}}(y_t^{t+k})) \right) \\
&= \left(H(\rho(y_t^{t+k+1})) - H(\rho_{\not x_t^{t+k}}(y_t^{t+k+1})) \right) - \sum_{i=1}^{k} [T_{x \to y}]_t^{t+i}
\end{aligned}
$$

Hence,

$$\sum_{i=1}^{k+1} [T_{x \to y}]_t^{t+i} = H(\rho(y_t^{t+k+1})) - H(\rho_{\mathscr{y}_t^{t+k}}(y_t^{t+k+1}))$$

Hence the proof.

$\square$

We provide following definition of average information transfer.

Definition 13. *[Average Information Transfer] We define the average information transfer from x to y as*

$$\bar{T}_{x \to y} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} [T_{x \to y}]_0^i \tag{3.23}$$

Lemma 14. *We have*

$$\bar{T}_{x \to y} = \bar{H}(Y) - \bar{H}_{\mathscr{x}}(Y)$$

where $\bar{H}(Y) = \lim_{n \to \infty} \frac{1}{n} H(\rho(y_0^n))$ is the entropy rate of y and $\bar{H}_{\mathscr{x}}(Y) = \lim_{n \to \infty} \frac{1}{n} H(\rho_{\mathscr{x}}(y_0^n))$ is the entropy rate of y, when x is held frozen.

Proof. The proof follows directly from the definition of average information and theorem (12).

$\square$

Though our concept of information transfer is way different than directed information, we will see later (chapter 4) that the average information transfer is indeed related to the average directed information.

3.3 Information Transfer and Dynamics of Linear Systems

A linear system may not exhibit the various complex dynamics that a nonlinear system can have, but still, the time domain trajectories of a linear system show different behaviours. The different behaviours that a linear system exhibit is governed by the eigenvalues of the system matrix. For example, if the eigenvalues of a linear discrete time system are outside the unit circle in the complex plane, then the trajectories diverge to infinity. Again, if they are on the unit circle, then the trajectories

oscillate. Similarly, if the eigenvalues are inside the unit circle, then the trajectories are stable, but depending on whether the eigenvalues are complex or real, the trajectories oscillate before settling down or does not oscillate. So one may think that the transient information transfer between the states should also show similar trends, depending on the location of the eigenvalues. In this section, we show, via simulations, that the transient one step information transfer in a linear system does show similar trends.

Case I : System is stable with real eigenvalues.

We consider a two dimensional linear system

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} 0 & 0.4 \\ 0.4 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

In this case, the eigenvalues are real and are inside the unit circle. Hence, the trajectories does not show any oscillations and this is shown in fig. 3.3(a).
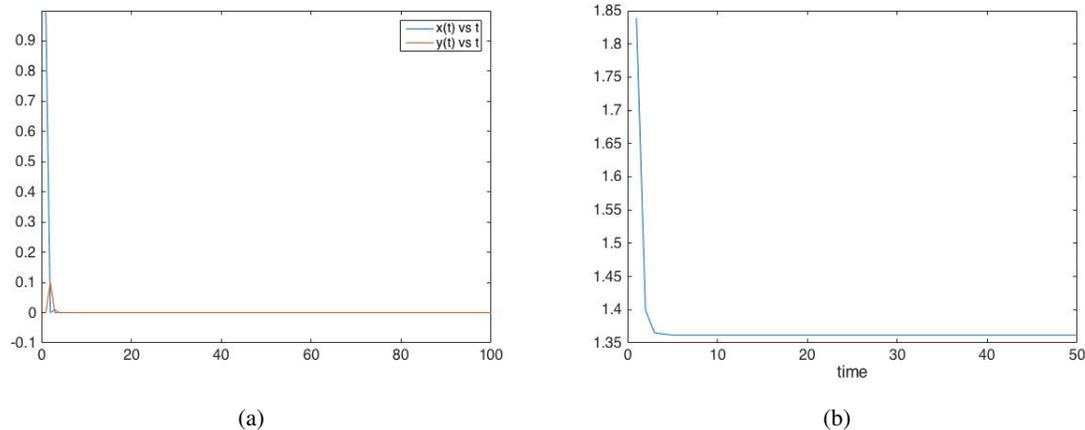
Figure 3.3   (a) Trajectories of the system. (b) IT from x to y.

The information transfer from x to y is shown in fig. 3.3(b). From the trajectory plot, we see that the trajectories show no oscillations and as expected, the transient information transfer also show no oscillation in the transients.

Case II : System is stable with complex eigenvalues.

We consider a two dimensional linear system

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -0.95^2 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

In this case, the eigenvalues are imaginary and are inside the unit circle. Hence, the trajectories initially show oscillations and then settle to the steady state value. This is shown in fig. 3.4(a).

(a)   (b)
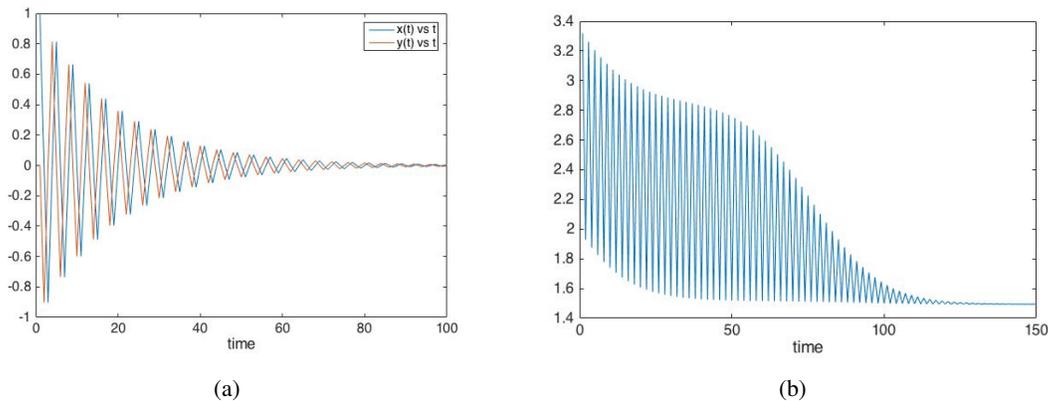
Figure 3.4   (a) Trajectories of the system. (b) IT from x to y.

The information transfer from x to y is shown in fig. 3.4(b). From the trajectory plot, we see that the trajectories show oscillations and as expected, the transient information transfer also shows oscillation in the transients and then settles to the steady state value.

Case III : System is oscillatory.

Next consider a two dimensional linear system

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

In this case, the eigenvalues are on the unit circle. Hence, the trajectories oscillates and this is shown in fig. 3.5(a).

From the trajectory plot, we see that the trajectories oscillate with constant amplitude and so does the transient information transfer, which is shown in fig. 3.5(b).

Case IV : System is unstable.
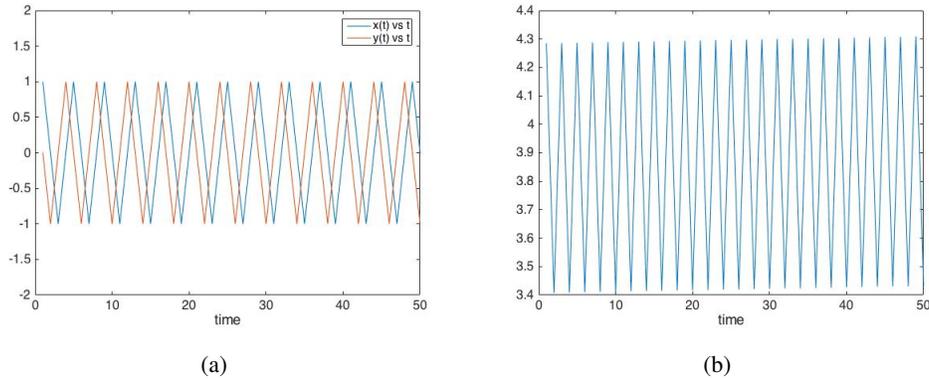
Figure 3.5 (a) Trajectories of the system. (b) IT from x to y.

We consider a two dimensional linear system

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1.01^2 & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

In this case, the eigenvalues are outside the unit circle and hence the system is unstable. The trajectories diverge to infinity and as the eigenvalues have imaginary part, the trajectories oscillate as well. This is shown in fig. 3.6(a) and the information transfer from x to y is shown in fig. 3.6(b).
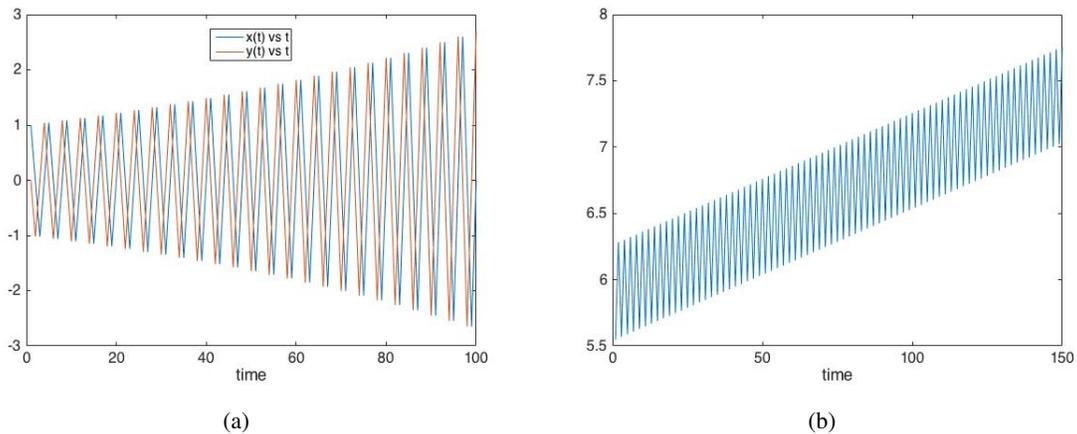
Figure 3.6 (a) Trajectories of the system. (b) IT from x to y.

In this case also, we find that the trajectories and the information transfer follow similar patterns. So it is not wrong to believe that the information transfer does throw light on the nature of dynamics

exhibited by a dynamical system. So one can use the information transfer to infer about the dynamics of a nonlinear system also and this is addressed in the next section.

3.4 Set Oriented Method for Computation of Information Transfer

The definition of information transfer between the states is based on the evolution of the probability density of the states and hence it is governed by the Perron-Frobenius operator. As we are going to see in a later chapter, in case of linear systems with additive Gaussian noise, the linearity property allows us to derive analytic expressions for the information transfer between the states. But for general discrete time maps, it is not possible to derive analytic closed form expressions for the information transfer. In fact, even in absence of noise, often it is difficult to derive an expression for the Perron-Frobenius operator and hence it will be difficult to obtain the information transfer. So, one has to look for numerical methods to compute the information transfer. In this section, we will describe set oriented numerical methods for finding the information transfer.

In order to obtain a finite-dimensional approximation of the continuous P-F operator, one considers a finite partition of the state space X as follows :

$$\mathcal{X} = \{D_1, D_2, \ldots, D_L\} \tag{3.24}$$

such that $X = \cup_{i=1}^{L} D_i$ and $D_i \cap D_j = \phi$ for all $i \neq j$. With this partition, the finite dimensional approximation of the P-F semi-group is a $L \times L$ matrix P such that

$$[P]_{ij} = \frac{m(f^{-1}(D_j) \cap D_i)}{m(D_i)} \tag{3.25}$$

where m is the Lebesgue measure. The ij^{th} entry of the finite dimensional P matrix gives the probability of transition from cell D_i to cell D_j. Hence $[P]_{ij} = \text{Prob}(z_{t+1} \in D_j | z_t \in D_i)$. Moreover, if $\mu_z = \begin{pmatrix} \mu_{z_1} & \mu_{z_2} & \cdots & \mu_{z_L} \end{pmatrix}$ is the density of z at time t, then

$$\mu_{z'} = \mu_z P \tag{3.26}$$

is the density of z at time step $(t+1)$.

In order to find the information transfer between the states in a system, we note that the one step information transfer from state x to state y can be written as

$$[T_{x \to y}]_t^{t+1} = H(\rho(y', y)) - H_{\not{x}}(\rho(y', y)). \tag{3.27}$$

This is because $H(\rho(y'|y)) = H(\rho(y', y)) - H(\rho(y))$ and $H(\rho(y)) = H_{\not{x}}(\rho(y))$. We will use equation (3.27) to compute the one step information transfer numerically.

Note that, the marginal distribution of (y', y) can be obtained if one can compute the joint distribution $\rho(z', z)$. Let μ_z be the distribution of z at time step t. Hence,

$$\text{Prob}(z' \in D_j, z \in D_i) = \text{Prob}(z \in D_i).\text{Prob}(z' \in D_j | z \in D_i) = \mu_z^i [P]_{ij} \tag{3.28}$$

So starting with an initial distribution μ_0, using (3.26) and (3.28) one can calculate the joint probability density of (z', z), as the system evolves in time and thus one can calculate the marginal distribution $\rho(y'y)$. Let $\mu_{y'y}$ be the distribution of the variable (y', y), then

$$H(\rho(y', y)) = -\sum [\mu_{y'y}]_i \log[\mu_{y'y}]_i \tag{3.29}$$

Similarly, one can compute the entropy of (y', y) when x is held frozen and thus obtain the information transfer. Next, we illustrate this method of computation of information transfer in two different systems, namely a linear system, where the information transfer can also be calculated analytically and a non-linear discrete time map.

Example 15 (Linear system)**.** *Consider the linear system*

$$\begin{aligned} x_{t+1} &= .9x_t + \xi_x \\ y_{t+1} &= 10x + .5y_t + \xi_y \end{aligned} \tag{3.30}$$

From the system equations, we see that while x enters the dynamics of y, y does not affect the dynamics of x and hence the information transfer from x to y should be non-zero, whereas, the transfer from y to x should be zero. In fact, in a later chapter, we prove the result for zero transfer for linear systems. Figure 3.7(a) shows the information transfer from state x to state y, calculated using the set-oriented method and also using the formula for information transfer in a linear system. We find that though the values in the transients are not that close, as the system reaches steady state, the values of the
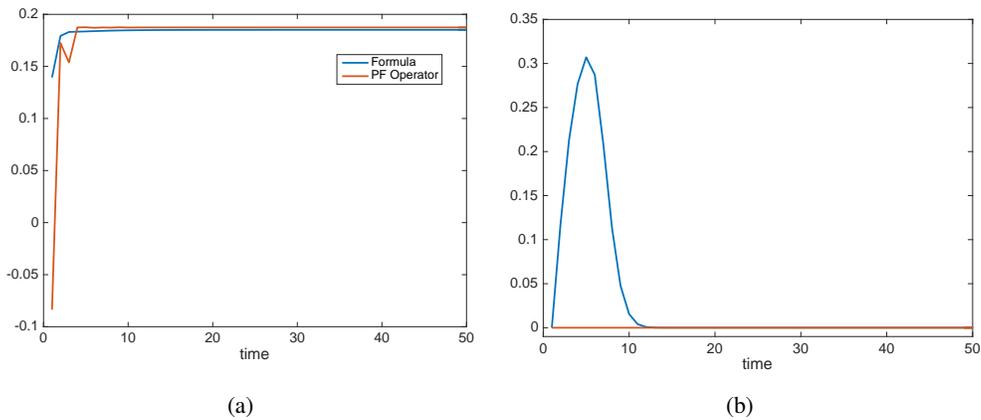
Figure 3.7   (a) IT from x to y. (b) IT from y to x.

transfer match closely. Fig. 3.7(b) shows the information transfer from y to x. From the theory (we prove this result later) we know that the information transfer from y to x should be zero. This is shown by the plot obtained by using the formula for information transfer. The other plot shows the transfer obtained using the set-oriented methods and we find that thought the transfer is non-zero in the initial stages, it goes to zero as the system evolves and thus validating the use of set-oriented methods to calculate the information transfer. The value of the transfer, calculated using the set-oriented method depends on a number of factors like how the state space is discretized and how many initial points are chosen to obtain the finite dimensional P-F matrix. However, as the system evolves, these factors play less and less important roles and as the system approaches the steady state, the values of the transfer calculated using the set-oriented method approached the value obtained using the analytic expression of transfer.

Example 16 (Henon map)**.** *In this example we consider the Henon map which is given by*

$$x' = 1 + y - ax^2$$
$$y' = bx$$

$$(3.31)$$

with $a, b > 0$.

For the first set of simulations, we choose $a = 1.4$ and $b = .3$. This is when the Henon map exhibits chaotic behaviour and has a chaotic attractor, as shown in fig. 3.8(a). The corresponding invariant
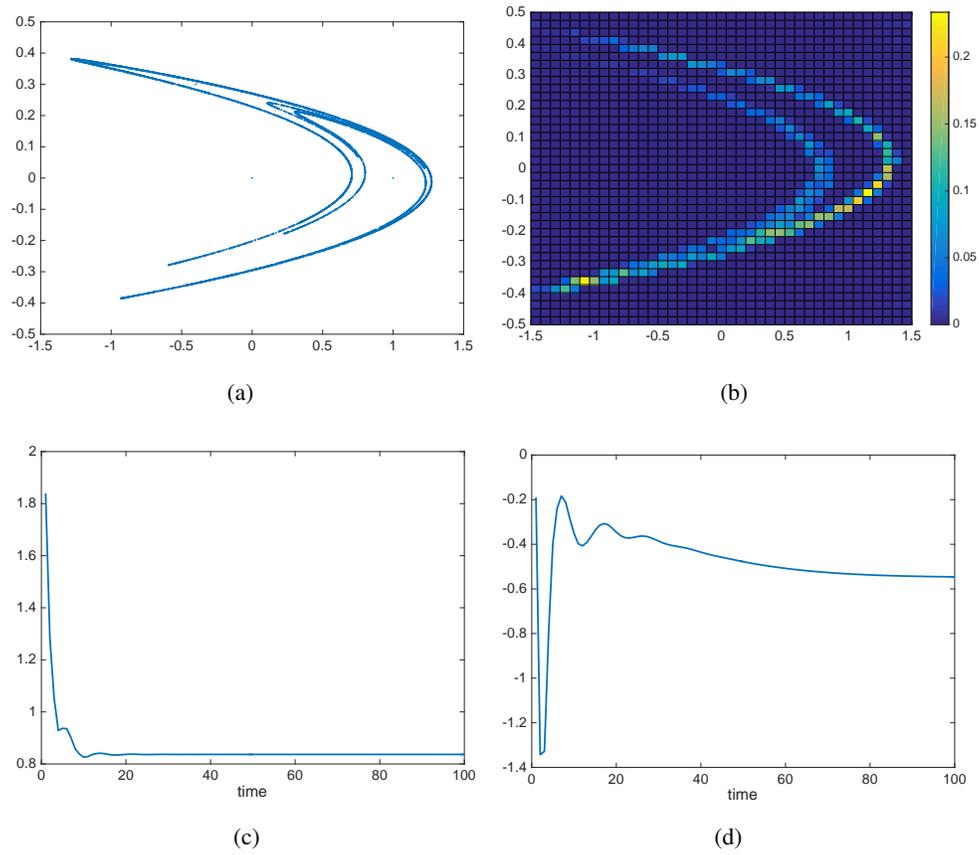
Figure 3.8 (a) Phase space trajectories (b) Invariant measure. (c) Information transfer from x to y. (d) Information transfer from y to x.

measure for the chosen set of parameters is shown in fig 3.8(b). The corresponding information transfers are shown in fig 3.8(c) and fig 3.8(d).
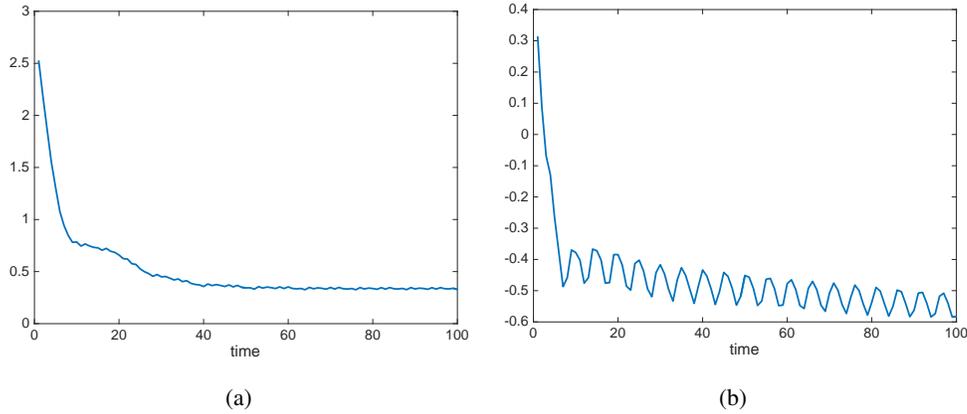
Figure 3.9 (a) Information transfer from x to y. (b) Information transfer from y to x.

For the second set of simulations simulation purposes, we choose $a = .18$ and $b = .7$. With these set of parameter values, the Henon map has a periodic orbit. Figure 3.9(a) shows the information transfer from x to y and fig. 3.9(b) shows the information transfer from y to x. Before reaching the steady state, from the information transfer plots, it is clear that x has a greater influence on y than what y has on x. Moreover, the information transfer oscillates and this is because of the presence of the periodic orbit.

Example 17 (Van der Pol Oscillator)**.** *Van der Pol oscillator is a non-conservative system with a non-linear damping and it evolves in time according to the second order differential equation*

$$\ddot{x} - \mu(1 - x^2)\dot{x} + x = 0 \tag{3.32}$$

where x is the position and μ is a scalar indicating the nonlinearity and the strength of damping. Using the variable $y = \dot{x}$, the system equations can be written as

$$\dot{x} = y$$

$$\dot{y} = \mu(1 - x^2)y - x$$

When $\mu = 0$, there is no damping and so the equation of motion becomes

$$\ddot{x} + x = 0$$

Hence, in this case, it represents a simple harmonic oscillator and there is conservation of energy.
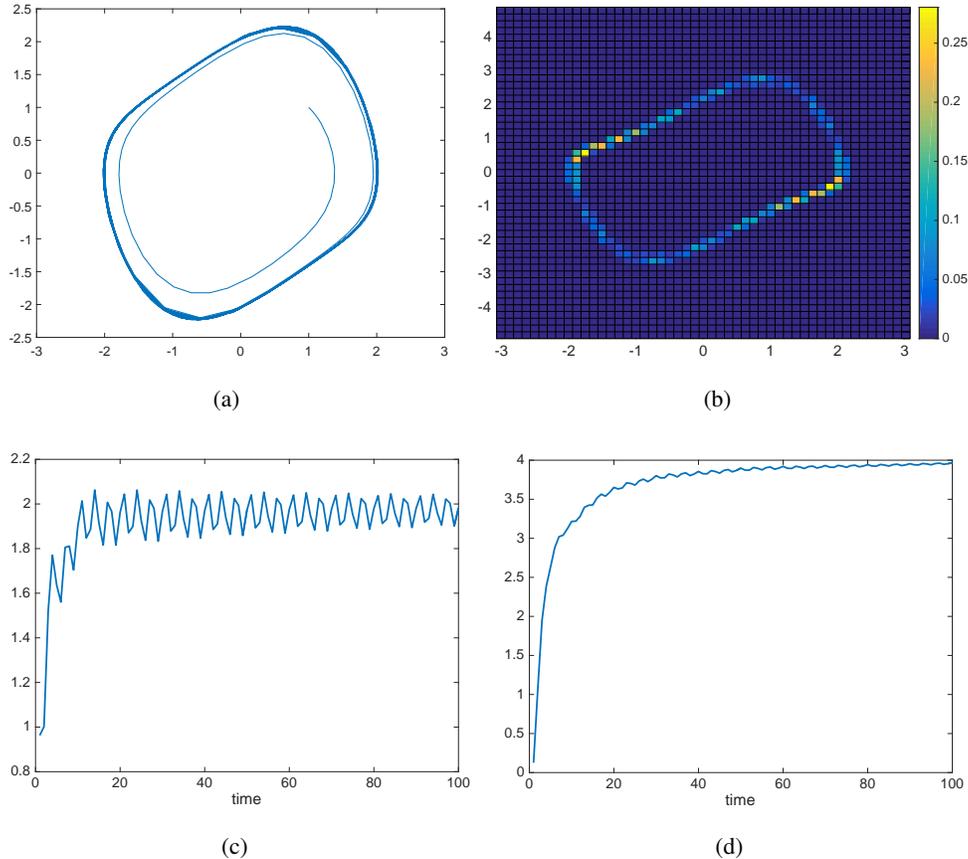
Figure 3.10   (a) Limit cycle for $\mu = 0.5$. (b) Corresponding invariant density. (c) Information transfer from x to y. (d). Information transfer from y to x.

When $\mu > 0$, the system has a limit cycle and for $\mu = 0.5$ the this is shown in fig *3.10*(a). The corresponding invariant measure is shown in fig. *3.10*(b). The information transfers from x to y and from y to x are shown in figures *3.10*(c) and *3.10*(d) respectively. We find that both the information transfers show oscillations which is because of the presence of the limit cycle. In fact, both in the Henon map example and the Van der Pol oscillator, we find that whenever there is a limit cycle in the system, the information transfers show oscillations.

In future, we hope to analyze complex dynamics from information transfer viewpoint and see how the information transfer throws light on the dynamics of the system. One starting point is the limit cycle behaviour, where we saw that the presence of limit cycle reflects in the information transfer. We

hope to capture more characteristics of complex systems in the future, using our proposed method of information transfer.

CHAPTER 4. INFORMATION TRANSFER IN LINEAR CONTROL SYSTEM

Information transfer is expressed as a difference of two entropies. As such, to compute it, knowledge about the evolution of densities, as a dynamical system evolves in time is necessary. For general nonlinear systems, there is no closed form expression for the evolution of densities. However, for linear dynamical systems, with a Gaussian assumption on the initial distribution of states and noise, one can derive closed-form expressions for information transfer.

4.1 Information Transfer Between the States

For simplicity, in this section we will consider a stochastic perturbed linear dynamical system without input. Let the system be given by

$$z(t+1) = Az(t) + \sigma\xi(t) \tag{4.1}$$

where $z(t) \in \mathbb{R}^N$ and $\xi(t)$ is vector valued Gaussian random variable with zero mean and unit variance. We assume that the initial conditions are Gaussian distributed with probability density function

$$\rho_0(z) = \frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma(0)|^{\frac{1}{2}}} \exp\left(z^\top \Sigma(0)^{-1} z\right)$$

Since the system is linear, the distribution of the system state for all future time will remain Gaussian with covariance $\Sigma(t)$ satisfying

$$A\Sigma(t-1)A^\top + \sigma^2 I = \Sigma(t)$$

To define the information transfer between various subspace we again introduce following notation to split the A matrix

$$z(t+1) = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} A_x & A_{xy} \\ A_{yx} & A_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \sigma\xi \tag{4.2}$$

The A matrix can be further split using the subspace decomposition $x = (x_1^\top, x_2^\top)^\top$ as follows:

$$\begin{pmatrix} A_x & A_{xy} \\ A_{yx} & A_y \end{pmatrix} = \begin{pmatrix} A_{x_1} & A_{x_1 x_2} & A_{x_1 y} \\ A_{x_2 x_1} & A_{x_2} & A_{x_2 y} \\ A_{y x_1} & A_{y x_2} & A_y \end{pmatrix} \tag{4.3}$$

Based on the decomposition of the system A matrix we can also decompose the covariance matrix Σ at time instant t as follows.

$$\Sigma = \begin{pmatrix} \Sigma_x & \Sigma_{xy} \\ \Sigma_{xy}^\top & \Sigma_y \end{pmatrix} = \begin{pmatrix} \Sigma_{x_1} & \Sigma_{x_1 x_2} & \Sigma_{x_1 y} \\ \Sigma_{x_1 x_2}^\top & \Sigma_{x_2} & \Sigma_{x_2 y} \\ \Sigma_{x_1 y}^\top & \Sigma_{x_2 y}^\top & \Sigma_y \end{pmatrix}$$

$$\tag{4.4}$$

Using the above notation, we state following theorem providing explicit expression for information transfer in linear dynamical system during transient and steady state.

Theorem 18. *Consider the linear dynamical system (4.1) and associated splitting of state space in Eqs. (4.2) and (4.3). We have following expression for information transfer between various subspace*

$$[T_{x \to y}]_t^{t+1} = \frac{1}{2} \log \frac{|A_{yx} \Sigma_y^s(t) A_{yx}^\top + \sigma^2 I|}{\sigma^2} \tag{4.5}$$

where $\Sigma_y^s(t) = \Sigma_x(t) - \Sigma_{xy}(t)\Sigma_y(t)^{-1}\Sigma_{xy}(t)^\top$ is the Schur complement of $\Sigma_y(t)$ in the matrix $\Sigma(t)$.

$$[T_{x_1 \to y}]_t^{t+1} = \frac{1}{2} \log \frac{|A_{yx} \Sigma_y^s(t) A_{yx}^\top + \sigma^2 I|}{|A_{yx_2}(\Sigma_y^s)_{yx_2}(t) A_{yx_2}^\top + \sigma^2 I|} \tag{4.6}$$

where $(\Sigma_y^s)_{yx_2}$ is the Schur complement of Σ_y in the matrix

$$\begin{pmatrix} \Sigma_{x_2} & \Sigma_{x_2 y} \\ \Sigma_{x_2 y}^\top & \Sigma_y \end{pmatrix}$$

Proof. From the formula of information transfer we have

$$[T_{x \to y}]_t^{t+1} = H(\rho(y'|y)) - H_{\not{x}}(\rho(y'|y))$$

So to compute the transfer, we need to know $\rho(y'|y) = \frac{\rho(y',y)}{\rho(y)}$ and $\rho_{\not{x}}(y'|y)$. Furthermore, since all the probability density function involved are Gaussian, we use following formula for the entropy of Gaussian density function.

$$H(\rho(w)) = \frac{1}{2}\log((2\pi e)^N |\Sigma(t)|)$$

for $\rho(w)$ of the form

$$\rho(w) = \frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma(t)|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}w^\top \Sigma(t)^{-1}z\right)$$

Hence to compute the required transfer, we only need to know the covariance of $\rho(y'|y)$ and $\rho_{\not{x}}(y'|y)$. The covariance matrix for the joint z' and z is given by

$$R_{z'z} = \begin{pmatrix} \Sigma_{z'} & \Sigma_{z'z} \\ \Sigma_{zz'} & \Sigma_z \end{pmatrix}, \quad \Sigma_z = \begin{pmatrix} \Sigma_x & \Sigma_{xy} \\ \Sigma_{xy}^\top & \Sigma_y \end{pmatrix}$$

where $\Sigma_{z'} = A\Sigma_z A^\top + \sigma^2 I$, $\Sigma_{zz'} = \Sigma_z A^\top$, $\Sigma_{z'z} = A\Sigma_z = \Sigma_{zz'}^\top$ Hence, we have

$$R_{y'y} = \begin{pmatrix} \Sigma_{y'} & \Sigma_{y'y} \\ \Sigma_{y'y}^\top & \Sigma_y \end{pmatrix}, y'|y \sim \mathcal{N}(-, \Sigma_{y'} - \Sigma_{y'y}\Sigma_y^{-1}\Sigma_{y'y}^\top)$$

$$\Sigma_{y'} = A_{yx}\Sigma_x A_{yx}^\top + A_y \Sigma_{xy} A_{yx}^\top + A_{yx}\Sigma_{xy}A_y^\top$$
$$+ A_y\Sigma_y A_y^\top + \sigma^2 I$$

$$\Sigma_{y'y} = A_{yx}\Sigma_{xy} + A_y\Sigma_y$$

Note that since the entropy is function of covariance, we ignore the explicit computation of mean for the underlying Gaussian random variable. After simplification, we obtain

$$\rho(y'|y) \sim \mathcal{N}(-, A_{yx}(\Sigma_x - \Sigma_{xy}\Sigma_y^{-1}\Sigma_{xy}^\top)A_{yx}^\top + \sigma^2 I).$$

For $\rho_{\not{x}}(y'|y)$ we look at the system, $y' = A_y y + \sigma\xi_y$. Hence, we have $\rho_{\not{x}}(y'|y) \sim \mathcal{N}\left(-, \sigma^2 I\right)$. Hence we have

$$T_{x\to y}(t) = \frac{1}{2}\log\frac{|A_{yx}\Sigma_y^s(t)A_{yx}^\top + \sigma^2 I|}{\sigma^2}$$

where $\Sigma_y^s(t) = \Sigma_x(t) - \Sigma_{xy}(t)\Sigma_y(t)^{-1}\Sigma_{xy}(t)^\top$.

To compute the information transfer from $x_1 \to y$, we need to compute covariance for $\rho_{\not x_1}(y^{'}|y)$. Hence, we look at the system where x_1 is absent. Hence, we look at

$$x_2^{'} = A_{x_2}x_2 + A_{x_2y}y + \xi_{x_2}$$

$$y^{'} = A_{yx_2}x_2 + A_y y + \xi_y \tag{4.7}$$

The expression for $\rho(y^{'}|y)$ remains unchanged from the previous case. To compute $\rho_{\not x_1}(y^{'}|y)$, we use system equation (4.7) and use procedure similar to one used in the derivation of $\rho_{\not x}(y^{'}|y)$ with x replaced with x_1. Hence, we have

$$\rho_{\not x_1}(y^{'}|y) = \mathcal{N}\left(-, A_{yx_2}(\Sigma_y^s)_{yx_2}(t)A_{yx_2}^{\top} + \sigma^2 I\right) \tag{4.8}$$

where,

$$(\Sigma_y^s)_{yx_2}(t) = (\Sigma_{x_2}(t) - \Sigma_{x_2y}(t)\Sigma_y^{-1}(t)\Sigma_{x_2y}^{\top}(t)).$$

Hence,

$$[T_{x_1 \to y}]_t^{t+1} = \frac{1}{2}\log\frac{|A_{yx}\Sigma_y^s(t)A_{yx}^{\top} + \sigma^2 I|}{|A_{yx_2}(\Sigma_y^s)_{yx_2}(t)A_{yx_2}^{\top} + \sigma^2 I|} \tag{4.9}$$

$\square$

The results from the above theorem can be used to provide general expression for information transfer between scalar state z_i and z_j for linear network system. In particular with no loss of generality we can assume $z_i = z_1$ and $z_j = z_2$, then the expression for $T_{z_1 \to z_2}$ can be obtained from (4.6) by defining

$$x_1 := z_1, \quad y = z_2, \quad x_2 := (z_3, \dots, z_N)$$

$$z_{\not y} = (z_1, z_3, z_4 \dots, z_N), z_{\not y\not y} = (z_3, z_4, \dots, z_N)$$

For linear systems with Gaussian noise, the one step zero transfer can be characterized by looking at system matrix A. In particular, we have the following theorem.

Theorem 19. $A_{z_j z_i} = 0$, if and only if $[T_{z_i \to z_j}]_t^{t+1} = 0$.

Proof. Without loss of generality, we prove the result for $i = 1$ and $j = 2$. With the notation

$$x_1 := z_1, \quad y = z_2,, \quad x_2 := (z_3, \ldots, z_N)$$

we have

$$[T_{x_1 \to y}]_t^{t+1} = \frac{1}{2} \log \frac{|A_{yx} \Sigma_y^s(t) A_{yx}^\top + \sigma^2 I|}{|A_{yx_2} (\Sigma_y^s)_{x_\gamma}(t) A_{yx_2}^\top + \sigma^2 I|} \tag{4.10}$$

where $A_{yx} = [A_{yx_1} \quad A_{yx_2}]$ and

$$\Sigma_y^s(t) =$$

$$\begin{bmatrix} \Sigma_{x_1} - \Sigma_{x_1 y} \Sigma_y^{-1} \Sigma_{x_1 y}^\top & \Sigma_{x_1 x_2} - \Sigma_{x_1 y} \Sigma_y^{-1} \Sigma_{x_2 y}^\top \\ \Sigma_{x_1 x_2}^\top - \Sigma_{x_2 y} \Sigma_y^{-1} \Sigma_{x_1 y}^\top & \Sigma_{x_2} - \Sigma_{x_2 y} \Sigma_y^{-1} \Sigma_{x_2 y}^\top \end{bmatrix}$$

So, if $A_{yx_1} = 0$,

$$A_{yx} \Sigma_y^s(t) A_{yx}^\top + \sigma^2 I$$

$$= A_{yx_2} (\Sigma_{x_2} - \Sigma_{x_2 y} \Sigma_y^{-1} \Sigma_{x_2 y}^\top) A_{yx_2}^\top + \sigma^2 I$$

$$= A_{yx_2} (\Sigma_y^s)_{x_\gamma}(t) A_{yx_2}^\top + \sigma^2 I$$

Hence,

$$\begin{aligned} [T_{x_1 \to y}]_t^{t+1} &= \frac{1}{2} \log \frac{|A_{yx} \Sigma_y^s(t) A_{yx}^\top + \sigma^2 I|}{|A_{yx_2} (\Sigma_y^s)_{x_\gamma}(t) A_{yx_2}^\top + \sigma^2 I|} \\ &= \frac{1}{2} \log \frac{|A_{yx_2} (\Sigma_y^s)_{x_\gamma}(t) A_{yx_2}^\top + \sigma^2 I|}{|A_{yx_2} (\Sigma_y^s)_{x_\gamma}(t) A_{yx_2}^\top + \sigma^2 I|} \\ &= 0 \end{aligned}$$

The other direction follows directly from the information transfer formula. $\qquad \square$

4.1.1 Examples

In this subsection, we look at two different simple examples to demonstrate how our formalism of information transfer can be used to infer about the causality structure and influence in a dynamical system.

Example 1 : Information transfer and causality. We consider a simple series RLC circuit with the

states as the voltage across the capacitor and the current. The state space equations, with input voltage to be zero, and noise $\xi = [\xi_1 \quad \xi_2]^\top$ are

$$
\begin{bmatrix} \dot{v}_c \\ \dot{i} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{C} \\ -\frac{1}{L} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} v_c \\ i \end{bmatrix} + \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}
\tag{4.11}
$$

We assume the initial voltage across the capacitor is $v_c(0)$. The resistor, capacitor and inductor

(a)   (b)
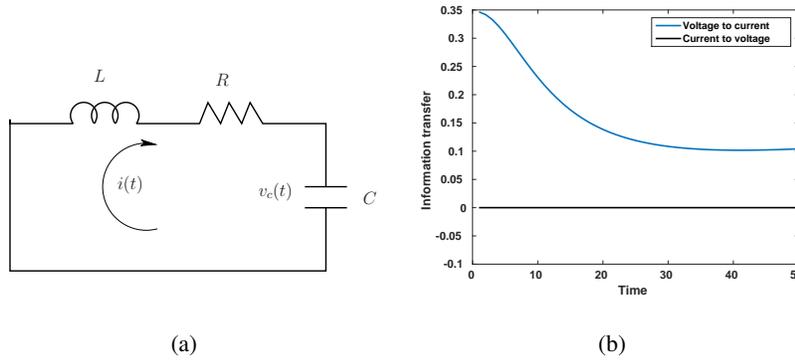
Figure 4.1   a) RLC series circuit; b) Information transfer between the states

values are 1 ohm, 1 farad and 1 Henry respectively. The circuit is shown in figure 4.1 (a). We discretize the system and look at the information transfer between the states for the discretized system. The information transfer is shown in figure 4.1 (b). We know that it is the voltage v_c that causes the current in the circuit. So, the voltage is the cause and the current is the effect. This can be inferred from the information transfer between the states. As we see, the information transfer from the voltage to the current is non-zero, whereas, the transfer from the current to the voltage is zero. Hence, the voltage affects the current, but the current does not affect the voltage. Hence, we infer that the voltage is the cause and current is the effect.

Example 2 : Information transfer and influence. In this example, we look at a mass spring system given in figure (5). We assume that $m \ll M$ and the spring constants are all equal (equal to one). Intuitively, since $m \ll M$, the mass m does not have much influence on the motion of the system, whereas, the bigger masses M has more influence on the motion of the system. This is again verified by the information transfer from the individual masses to the subspace formed by the other two masses. The information transfer from $M = 100$ to the rest of the system is 0.1121, while the transfer from

$m = 1$ to the rest of the system is 0.0095.
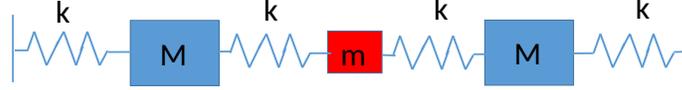
Figure 4.2   Mass spring system

So we say that the larger masses have more influence on the system compared to the smaller mass. This is also consistent with the concept of participation factor (28), where the authors have shown that the bigger masses *participate* a lot more than the smaller mass.

4.2   Information Transfer in Linear Control Systems

In this section, we derive an expression for information transfer in control dynamical system. In particular, we derive expressions for information transfer from state to output, input to state, and input to output. For the simplicity of discussion, we also restrict our discussion to single input single output case. We consider the following linear linear time invariant system with input and output.

$$z(t + 1) = Az(t) + Bu(t)$$
$$\theta(t) = Cz(t) + \omega(t)$$

(4.12)

where $z \in \mathbb{R}^N$ are the states, $A \in \mathbb{R}^{N \times N}$, B and C are of appropriate dimensions, $u(t) \in \mathbb{R}$ is the input, $\theta \in \mathbb{R}$ is output and $\omega(t)$ is the output noise, which is assumed to be zero mean i.i.d. Gaussian noise. For the simplicity of presentation, we will restrict our discussion to the case where the state space z is split in only two subspace i.e., $z = (x^\top, y^\top)^\top$ and the inputs and outputs are one dimensional. With this assumption, we have following splitting of the A, B, and C matrices.

$$A = \begin{pmatrix} A_x & A_{xy} \\ A_{yx} & A_y \end{pmatrix}, B = \begin{pmatrix} B_x \\ B_y \end{pmatrix}, C = \begin{pmatrix} C_x & C_y \end{pmatrix}$$

(4.13)

4.2.1 Information from Input to State

The evolution of the state x is $x' = A_x x + A_{xy} y + B_x u$. In deriving the formulas in this subsection, we think of the input $u(t)$ as a i.i.d. Gaussian variable such that

$$\mathrm{E}[u(t)u(t+k)^\top] = 0, \forall k \neq t \text{ and } \mathrm{E}[zu^\top] = 0.$$

Hence,

$$\Sigma_{x'|x} = A_{xy}\Sigma_x^s A_{xy}^\top + B_x \Sigma_u B_x^\top \tag{4.14}$$

where Σ_x^s is the Schur complement of Σ_x in Σ_z and Σ_u is the covariance of the input at time t.

When the input u is held frozen,

$$\Sigma_{x'|x}^{\not u} = A_{xy}\Sigma_x^s A_{xy}^\top \tag{4.15}$$

Hence the information transfer from u to x is

$$T_{u \to x} = \frac{1}{2} \log \frac{|A_{xy}\Sigma_x^s A_{xy}^\top + B_x \Sigma_u B_x^\top|}{|A_{xy}\Sigma_x^s A_{xy}^\top|} \tag{4.16}$$

where $|\cdot|$ is the determinant.

4.2.2 Information from State to Output

In this subsection, we look at the information transfer from the states of the system to the output of the control system. For simplicity, we only look at the information transfer from the entire state space z to the output θ. The general case of the information transfer from any state z_i to any output θ_j for a MIMO system can be dealt with similarly.

From the output equation, we have

$$\Sigma_{\theta'|\theta} = C\Sigma_{z'}C^\top + \Sigma_\omega$$
$$-(CA\Sigma_z C^\top)(C\Sigma_z C^\top)^{-1}(CA\Sigma_z C^\top)^\top$$

where $\Sigma_{z'} = A\Sigma_z A^\top + B\Sigma_u B^\top$ and Σ_ω is the covariance of the output noise.

When z is frozen, the output equation is

$$\theta_{\not z}(t) = \omega(t) \tag{4.17}$$

Hence, $\Sigma_{\theta'|\theta}^{\not{z}} = \Sigma_\omega$.

So the transfer is

$$T_{z \to \theta} = \frac{1}{2} \log \frac{|\Sigma_{\theta'|\theta}|}{|\Sigma_\omega|}$$

where

$$\Sigma_{\theta'|\theta} = C\Sigma_{z'}C^\top + \Sigma_\omega$$
$$-(CA\Sigma_z C^\top)(C\Sigma_z C^\top)^{-1}(CA\Sigma_z C^\top)^\top$$

4.2.3 Information from Input to Output

As before, we address the case of SISO systems for simplicity. We have

$$\Sigma'_\theta = CA\Sigma_z A^\top C^\top + CB\Sigma_u B^\top C^\top$$

Hence,

$$\Sigma_{\theta'|\theta} = CA\Sigma_z A^\top C^\top + CB\Sigma_u B^\top c^\top + \Sigma_\omega$$
$$-(CA\Sigma_z C^\top)(C\Sigma_z C^\top)^{-1}(CA\Sigma_z C^\top)^\top \tag{4.18}$$

When u is frozen, we have

$$z' = Az$$
$$\theta = Cz + \omega \tag{4.19}$$

Hence,

$$\Sigma_{\theta'_{\not u}|\theta_{\not u}} = CA\Sigma_z A^\top C^\top + \Sigma_\omega$$
$$-(CA\Sigma_z C^\top)(C\Sigma_z C^\top)^{-1}(CA\Sigma_z C^\top)^\top \tag{4.20}$$

Hence the transfer is

$$T_{u \to \theta} = \frac{1}{2} \log \frac{|\Sigma_{\theta'|\theta}|}{|\Sigma_{\theta'_{\not u}|\theta_{\not u}}|} \tag{4.21}$$

where $\Sigma_{\theta'|\theta}$ and $\Sigma_{\theta'_{\not u}|\theta_{\not u}}$ are given by equations (4.18) and (4.20) respectively.

4.3   Information Transfer and Structural Controllability and Observability

Structural controllability was developed by (29), and is a weaker notion of controllability, where a system is said to structurally controllable, if either the system is controllable or is controllable when the non-zero entries of the system matrix are perturbed so that the perturbed system becomes controllable. Associated with a system

$$z(t + 1) = Az(t) + \xi(t), \quad z \in \mathbb{R}^N \tag{4.22}$$

is a directed graph $Z = (V_z, E_z)$ where $V_z = \{z_i | i = 1, \ldots, N\}$ is the node set and E_z is the edge set, such that there is directed edge from z_i to z_j iff ji^{th} entry of A is non-zero. Now if a single control input is placed at node k, then one can construct an extended graph $G = (V, E)$, with one extra node u and one extra edge from u to z_k. With this, the system is said to structurally controllable if all the nodes of the extended graph can be reached from node u (30). Since u is directly connected to z_k, one can say that the system, with input at z_k, is structurally controllable if there exist paths to all the nodes of Z from z_k.

In this section, we show that if information transfer from the input to all the nodes is non-zero, then the system is structurally controllable. Note that in this section, by information transfer, we will mean n-step information transfer for $n \in \mathbb{Z}_{>0}$.

Theorem 20. *If the k-step information transfer from state z_i to z_j of system (4.22) is non-zero for some $k \in \mathbb{Z}_{>0}$, then there is a directed path from node z_i to z_j in the corresponding directed graph for the system.*

Proof. We prove this by contrapositive argument. Without loss of generality, we assume $i = 2$ and $j = 1$. Suppose there is no path from z_2 to z_1. Hence, $[A^k]_{12} = 0$ for all $k \in \mathbb{Z}$ [1]. We show that in this case the k-step information transfer from z_2 to z_1 is zero for all $k \in \mathbb{Z}$. The evolution of the states can be written as

$$z(k) = A^k z(0) + A^{k-1}\xi(0) + \ldots + \xi(k - 1) \tag{4.23}$$

---

[1] Here $[A^k]_{ij}$ means the ij^{th} entry of A^k

Hence,

$$
\begin{aligned}
z_1(k) &= \sum_{j=1, j\neq 2}^{N} [A^k]_{(1j)} z_j(0) + \sum_{j=1, j\neq 2}^{N} [A^{k-1}]_{(1j)} \xi_j(0) \\
&\quad + \ldots + \xi_1(k)
\end{aligned}
\tag{4.24}
$$

since $[A^k]_{12} = 0$ and when z_2 is held frozen, we have

$$
\begin{aligned}
z_{1\not{2}}(k) &= \sum_{j=1, j\neq 2}^{N} [A_{\not{2}}^k]_{(1j)} z_j(0) + \sum_{j=1, j\neq 2}^{N} [A_{\not{2}}^{k-1}]_{(1j)} \xi_j(0) \\
&\quad + \ldots + \xi_1(k)
\end{aligned}
\tag{4.25}
$$

where $A_{\not{2}}$ is the system matrix when z_2 is held frozen, that is $A_{\not{2}}$ is obtained from A by deleting the second row and second column of A.

Now,

$$
\begin{aligned}
[A^2]_{(1j)} &= \sum_{i=1}^{N} a_{1i} a_{ij} = a_{12} a_{2j} + \sum_{i\neq 2} a_{1i} a_{ij} \\
&= \sum_{i\neq 2} a_{1i} a_{ij} = [A_{\not{2}}^2]_{1j}
\end{aligned}
\tag{4.26}
$$

since $a_{12} = 0$.

Similarly,

$$
\begin{aligned}
[A^3]_{1j} &= [A^2 A]_{1j} = [A_2 A]_{1j} = \sum_i (a_2)_{1i} a_{ij} = \\
(a_2)_{12} a_{2j} + \sum_{i\neq 2} (a_2)_{1i} a_{ij} &= \sum_{i\neq 2} (a_2)_{1i} a_{ij} = [A_{\not{2}}^3]_{1j}
\end{aligned}
$$

since $[A^2]_{12} = (a_2)_{12} = 0$. Let $[A^l]_{1j} = [A_{\not{2}}^l]_{1j}$ for some $l \in \mathbb{Z}_{>0}$. Then

$$
\begin{aligned}
[A^{l+1}]_{1j} &= [A^l A]_{1j} = [A_l A]_{1j} = \sum_i (a_l)_{1i} a_{ij} = \\
(a_l)_{12} a_{2j} + \sum_{i\neq 2} (a_l)_{1i} a_{ij} &= \sum_{i\neq 2} (a_l)_{1i} a_{ij} = [A_{\not{2}}^{l+1}]_{1j}
\end{aligned}
$$

Hence, $[A^k]_{1j} = [A_{\not{2}}^k]_{1j}$ for all $k \in \mathbb{Z}$. Hence from (4.24) and (4.25), $z_1(k) = z_{1\not{2}}(k)$. So,

$$
H(z_1(k)|z_1(k-1)\ldots z_1(0)) = H_{\not{2}}(z_1(k)|z_1(k-1)\ldots z_1(0)).
$$

Hence $[T_{z_2 \to z_1}]_0^k = 0$ for all $k \in \mathbb{Z}$. Hence, if there is non zero information transfer from z_2 to z_1, then there exists a directed path from z_2 to z_1. $\qquad \square$

Theorem 21. *If for some $n \in \mathbb{Z}_{>0}$, the information transfer from the input to all the states are non-zero, then the system is structurally controllable.*

Proof. We consider the extended graph associated with the dynamical system and consider the input node as an extra state. From theorem (20), if the information transfer from the input to any state z_i is non-zero, then there exists a path from the input node to the node z_i in the directed graph associated with the system. So, if the information transfer from the input to all the states is non zero, then there exists directed paths from the input node to all the other nodes in the directed graph associated with the system and hence the system is input reachable and hence is structurally controllable. $\qquad\square$

Results for structural observability follows from duality and can be stated as

Theorem 22. *If the information transfer from all the states to the output is non zero, then the system is structurally observable.*

Here we have proved the result for structural controllability for SISO case, but the result for MIMO case is similar.

4.4 Information Transfer in Feedback Control Systems

Information theory and feedback systems are interlinked and researchers have studied their interplay (31; 32; 33; 34). Again, Massey (11) had shown that the feedback capacity (C_{fb}) satisfies

$$C_{fb} \leq \lim_{n\to\infty} \max_{\rho(X^n||Y^{n-1})} \frac{1}{n} I(X^n \to Y^n) \tag{4.27}$$

where $\rho(X^n||Y^{n-1})$ is the conditional distribution of X^n, causally conditioned on Y^{n-1}. Again, in (35), it has been proved that in a feedback system, the Bode sensitivity transfer function from the output to the input is equal to the average directed information from the output (w) to the input (u), where the average directed information from the output to the input is defined as

$$\lim_{n\to\infty} \frac{1}{n} I(w^n \to u^n) \tag{4.28}$$

In this section, we consider the following feedback control system (figure 4.3)

$$z(t+1) = Az(t) + Bu(t)$$
$$w(t) = Cz(t), \quad u(t) = w(t) + \xi(t) \tag{4.29}$$
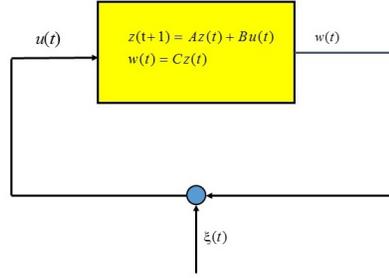
Figure 4.3   Feedback control system

and we are interested in computing information flow from output to input of the feedback control system i.e., $T_{w \to u}$. This is the information transfer from the plant to itself. Towards this goal we write the feedback control system as follows:

$$z(t+1) = (A + BC)z(t) + B\xi(t)$$

$$w(t) = Cz(t), \quad u(t) = w(t) + \xi(t) \tag{4.30}$$

Hence, $H_\psi(\rho(u'|u)) = H(\rho(\xi'))$. With this, and theorem (12), next we show that the Bode sensitivity transfer function, $\mathcal{S}$, from w to u is same as the average information transfer from w to u.

Theorem 23.

$$
\begin{aligned}
\bar{T}_{w \to u} &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \log |\mathcal{S}(e^{j2\pi\theta})| d\theta \\
&= \sum_{i=1}^{m} \log |\lambda_i(A_u)|
\end{aligned}
\tag{4.31}
$$

where $\lambda_i(A_u)$ are the unstable eigenvalues of the open loop system.

Proof. From theorem (12), we have

$$
\begin{aligned}
\bar{T}_{w \to u} &= \lim_{T \to \infty} \frac{1}{T} \sum_{i=1}^{T} (T_{w \to u})_0^i \\
&= \lim_{T \to \infty} \frac{1}{T}[H(u^T) - H_\psi(u^T)] = \lim_{T \to \infty} \frac{1}{T}[H(u^T) - H(\xi^T)] \\
&= \lim_{T \to \infty} \frac{1}{2T} \log \frac{|\Sigma_{u^T}|}{|\Sigma_{\xi^T}|} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \log \frac{|u(e^{j2\pi\theta})|^2}{|\xi(e^{j2\pi\theta})|^2} d\theta \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} \log |\mathcal{S}(e^{j2\pi\theta})| d\theta = \sum_{i=1}^{m} \log |\lambda_i(A_u)|
\end{aligned}
$$

where $\lambda_i(A_u)$ are the unstable poles of A.

$\square$

Hence, the average directed information and average information transfer are the same for the above feedback control system. This theorem, along with the fact that the channel capacity of a feedback channel is related to directed information, presents further evidence that our formulation of information transfer, though developed from a dynamical systems point of view, is consistent with already existing concepts in information theory and control theory. We feel this will allow us to study the concepts of information theory, control theory and dynamical systems from a more fundamental and common point of view.

Example 24. *Consider the feedback system given in figure 5 with*

$$A = \begin{pmatrix} 4 & 2 \\ 0 & 0.7 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad C = [\ -3.5 \ \ 0\ ]$$

We also assume that the noise ω is i.i.d. zero mean unit variance Gaussian white noise. The open
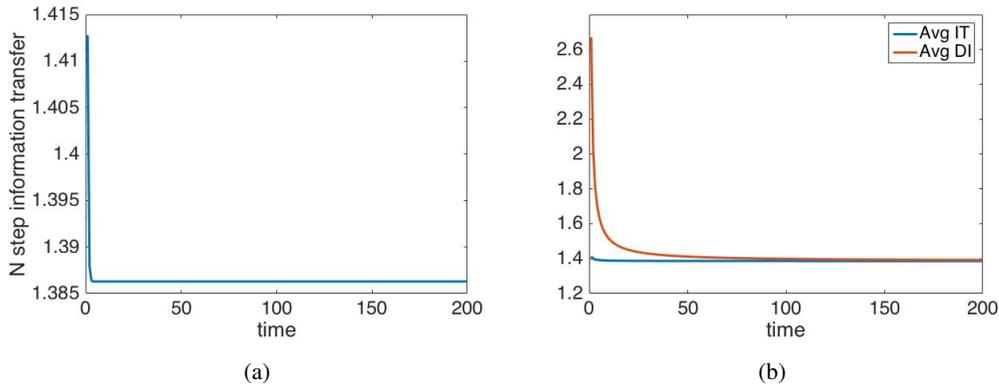
Figure 4.4 (a) N-step information transfer; (b) Average information transfer and average directed information

loop system has unstable pole at 4. Hence the Bode integral of the sensitivity transfer function is $\log(4) = 1.3863$. Figure 4.4(a) shows the N-step information transfer from the output of the system to the input of the system. The steady state value of the information transfer is 1.3863. Similarly, the average information transfer from the output to the input is also 1.3863. This is shown in figure

4.4(b). In this figure we also plot the average directed information transfer from the output to input. Here again we observe that the average information transfer converges to the Bode integral of the sensitivity transfer function much faster than the average directed information.

CHAPTER 5. INFORMATION TRANSFER AS A MEASURE OF CAUSALITY

5.1 Directed Information as a Measure of Causality

Directed information and transfer entropy are two of the most popular notions of information transfer used to characterize causality. In this section, we show that these notions of information transfer cannot faithfully capture the true causality structure in control dynamical system.

Directed information was first defined by Massey (11) as a generalization of Marko's bidirectional information (25). Both bidirectional information and directed information gave a sense of direction to Shannon's information theory and is viewed as a generalized information theory. Let $X^n = \{X_1, X_2, \cdots, X_n\}$ and $Y^n = \{Y_1, Y_2, \cdots, Y_n\}$ be two stochastic processes, viewed as a sequence of random variables. Massey and Kramer (12) defined the directed information from X^n to Y^n as

$$I(X^n \to Y^n) = H(Y^n) - H(Y^n \parallel X^n) \tag{5.1}$$

where $H(Y^n)$ is the entropy of the sequence Y^n and

$$H(Y^n \parallel X^n) := \sum_{i=1}^{n} H(Y_i | Y^{i-1}, X^i) \tag{5.2}$$

is the entropy of Y^n *causally* conditioned on X^n. This differs from

$$H(y^N | x^N) = \sum_{n=1}^{N} H(y_n | y^{n-1} x^N)$$

in the fact that x^n is replaced by x^N in each term of equation (1.3).

As for transfer entropy, it is like a Kullback entropy and in the case of a Markov chain of order one, Schreiber's formula for transfer entropy from x to y can be written as

$$T^S_{x \to y} = -\sum P(y^{n+1}, y^n) \log P(y^{n+1}|y^n) - \\ \left[-\sum P(y^{n+1}, y^n, x^n) \log P(y^{n+1}|y^n, x^n) \right] \tag{5.3}$$

In both directed information and transfer entropy, the conditioning is done on the entire past and so both these measures carry information about the entire past. So, if one likes to capture the one step information transfer, the most natural definition of information transfer from x to y would be

$$\hat{T}_{x \to y} = H(\rho(y'|y)) - H(\rho(y'|y, x)). \tag{5.4}$$

The directed information is asymmetric and gives a directional sense to the information and defines a measure to determine the direction of information flow. Since the development of the concept of directed information, this has been used in many different applications, like determining the channel capacity of a communication channel. Moreover, for Gaussian variables, directed information and Granger causality (8), (6) are equivalent (36). This allows the concept of directed information to be used as a measure of causality and hence it has been used to infer about the causal structure of statistical processes. In the following, we demonstrate using three different examples how the directed information fails to capture the true causality structure in dynamical systems setting. While the arguments are made in the context of directed information, similar conclusions can be drawn in the context of transfer entropy. We claim that at the heart of the problem is the manner in which causal conditioning is performed in both these definitions of information transfer.

5.1.1 Examples

Example 25. *Consider the following linear system with output*

$$\begin{pmatrix} x_1(t+1) \\ x_2(t+1) \\ x_3(t+1) \end{pmatrix} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \sigma \xi(k)$$

$$y(t) = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} x(t) \tag{5.5}$$

where $x_i(t)$ and $y(t)$ are the states and output respectively at time step t and $\xi(t)$ is a independent identically distributed (i.i.d.) zero mean unit variance Gaussian noise and σ is a constant. We notice
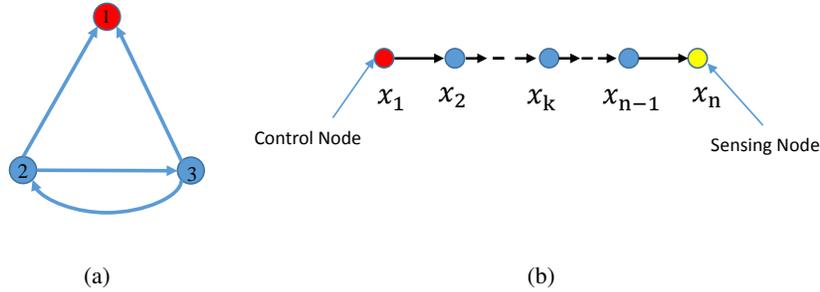
Figure 5.1   (a) Graph of dynamical system of example 1. (b) Graph of dynamical system of example 2.

that while there is directed path from $x_2 \to x_1$ and $x_3 \to x_1$ (fig. 5.1(a)), there is no path from x_1 to x_2 or from x_1 to x_3. Hence, we conclude that x_1 is not a cause of x_2 or x_3, that is, x_1 is not influencing x_2 or x_3. So if we treat the dynamical system as a stochastic process, we expect that the flow of information from $x_1 \to x_2$ and $x_1 \to x_3$ to be zero, thereby inferring that there is no causal connection from $x_1 \to x_2$ and $x_1 \to x_3$. In fact closer examination reveals that x_1 is not influencing x_2 and x_3 over any number of time steps. However, as we show below, the directed information fails to capture this true causal interaction between state x_1, x_2, and x_3. In particular, we show that $I(x_1^n \to x_2^n) \neq 0$ and $I(x_1^n \to x_3^n) \neq 0$ for any n. In Fig. 5.2(a), we plot the directed information from $x_1 \to x_2$ and $x_1 \to x_3$ over different time steps.
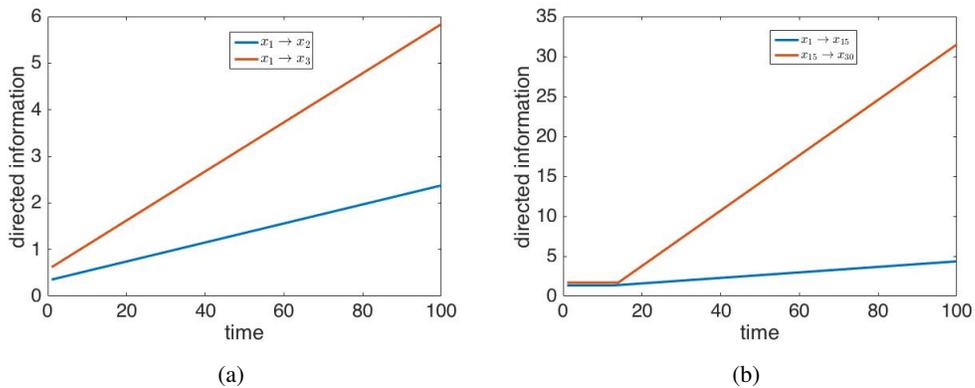
Figure 5.2   (a) Directed information plots for example 1. (b) Directed information plots for example 2.

Example 26. *Next we consider a single input single output linear system*

$$x(t+1) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} u$$

$$y(t) = \begin{pmatrix} 0 & \dots & 1 \end{pmatrix} x(t) \tag{5.6}$$

where $x(t) \in \mathbb{R}^N$. We note that the state x_1 is directly controlled through u and measurements are made at state x_N. State x_k, for $1 << k << N$, is assumed to be some intermittent state. Since the control input u does not affect state x_k instantaneously, but through series of k delays, we expect the information flow from state x_1 to x_k to be zero for time step $t = 0, \dots, k-1$. On the other hand since x_k influence x_N again through series of delay we expect the information flow from x_k to x_N to zero for $t = k, \dots, N-k-1$. However as we show in Fig. 5.2(b), directed information from $x_1 \to x_k$ and $x_k \to x_N$ is nonzero for all time $t = 0, \dots$, conveying that there is instantaneous flow of information from x_1 to x_k and from x_k to x_N. The simulation results in Fig. 5.2(b) are obtained for $N = 30$ nodes with $k = 15$. This example demonstrates that the directed information fails to capture indirect path of influence.

Example 27. *We again consider the system in Eq. (5.5) with measurement made at state x_2. We have already seen that the directed information flow computed over any time step from $x_1 \to x_2$ is nonzero. Naturally, question arise as to the significance of the information flow from $x_1 \to x_2$ for estimation problems. In fact there is extensive literature on the connection between directed information and Kalman filter (37). However, simple calculation show that the system matrix pair (A, C), from Eq. 5.5, is not observable with state x_1 being the only state not observable. In fact state x_1 is not only unobservable it is also structurally unobservable. Using similar argument it can be show while there is non zero flow of directed information from $x_1 \to x_2$ and $x_1 \to x_3$, the state x_2 and x_3 are not structurally controllable with control placed at state x_1.*

The above examples reveal some serious deficiencies of directed information as a measure of causality in dynamical system setting. Furthermore, the directed information value itself has no impli-

cation on two of the most fundamental system theoretic concepts of controllability and observability. This suggests that the directed information is a measure of *statistical* interaction between two signals but it does not successfully capture the true *dynamical* interactions between dynamical states. We expect that any meaningful definition of information transfer in network dynamical system should be able to capture true causal interaction and also obey time constraint of information flow.

5.2   Information Transfer Captures the True Causal Structure

In this section, we revisit the examples presented in section (5.1) and show that the new definition of information overcome some of the issues with directed information.
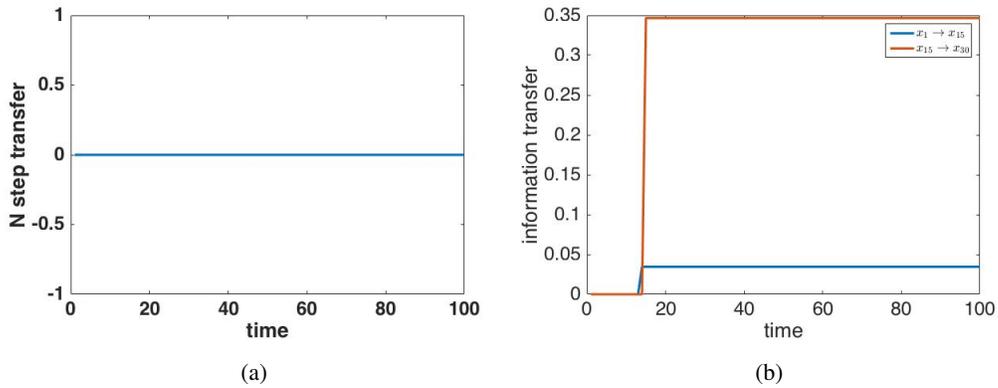
Figure 5.3   (a) N-step information transfer for example 1. (b) N-step information transfer for example 2.

Example 25 revisited : For the three state example discussed in section 5.1, we compute the information transfer using our proposed definition of information transfer and the formula (4.6). In Fig. 5.3(a), we plot the information transfer over n time steps and we notice that the information transfer from $x_1 \to x_2$ and $x_1 \to x_3$ are identically zero.

Example 26 revisited : We compute the information transfer from $x_1 \to x_k$ and from $x_k \to x_N$ for $k = 15$. Since, x_1 is connected to x_k through series of k delays, as expected the information transfer $T_{x_1 \to x_k}$ is zero over $n < k$ time step and for $n = k$ the information transfer jumps from zero to non-zero value as shown in Fig. 5.3(b). Similarly information transfer from $x_k \to x_N$ also shows a sudden jump. This proves that our proposed definition of information transfer obeys the time

constraints of transfer.

Example 27 revisited : In this example, since there is no flow of information from x_1 to x_2 or x_3, we conclude, from theorem (21) that the system is not structurally controllable from x_1. Similarly, the system is not structurally observable from x_2 or x_3 [theorem (22)].

CHAPTER 6. EXAMPLES AND APPLICATIONS

In this chapter, we look at two different applications of our proposed formalism of information transfer. One application is in power systems, where we study the contribution of individual generators to the inter-area oscillation modes of the power system. In the other example, we use the information transfer to characterize the influential nodes in a Twitter network. In the third example, we use the developed framework to find the optimal location of actuators in a double gyre fluid flow vector field.

6.1 Application to social network

Given a network, it is natural to ask which are the *important* nodes in the graph. For example, a social network plays a fundamental role as a medium for the spread of information, ideas, and influence among its members. An idea or innovation will appear and it can either die out quickly or make significant inroads into the population. If the idea spreads significantly in the network, it is natural to think that the idea originated in some influential node in the graph, whereas, if the idea dies down quickly, then one may think that the node of origin of the idea is not that influential in the network. In the context of social networks, one natural question is to determine links weight between two nodes that take into account the influence the two nodes exhibit on one another. We propose the use of the information transfer between the two nodes for assigning the links weights.

We are particularly interested in a directed social network such as Twitter or Google plus. Let A be the adjacency matrix consisting of $\{0, 1\}$ entries corresponding to the directed social network. If $A_{ij} = 1(0)$ then there is (no) directed path from node j to node i. For assigning weights on the graph, we construct $\mathcal{A}$ as follows:

$$\mathcal{A} = \alpha A$$

where α is suitable constant to ensure $|\lambda_{max}(\mathcal{A})| < 1$.

6.1.1 Simulation

In Fig. 6.1a we show the adjacency matrix for a 200 node Twitter follower network (for the purpose of visualization we only show part of the adjacency matrix with first 50 nodes). The data for this matrix is obtained from https://snap.stanford.edu/data/. This adjacency matrix, after the normalization procedure outlined above, is used to find the information transfer.
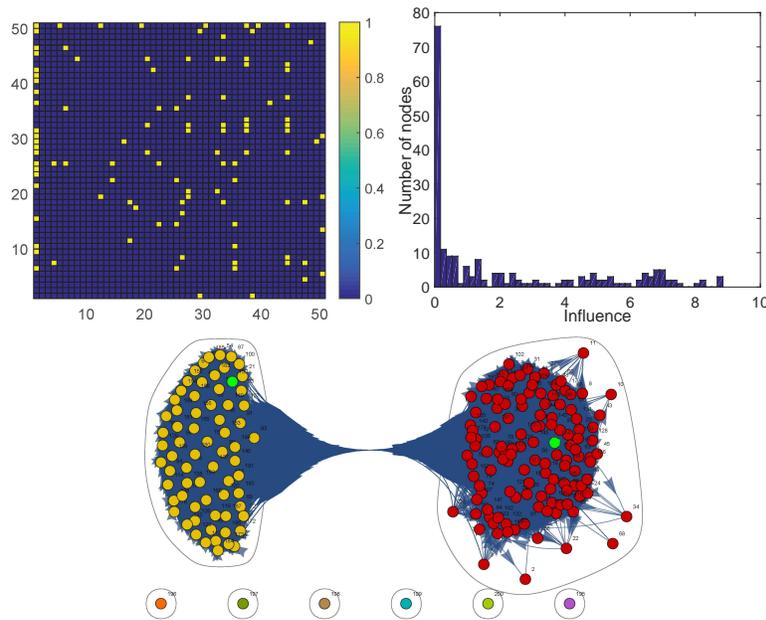
Figure 6.1   a) Twitter adjacency matrix. b) Distribution of influential nodes. c) Clustering based on influence in Twitter follower network.

In Fig. 6.1b, we plot the distribution of the influential nodes for the Twitter network. To obtain the influence distribution plot we did a histogram for the row sum of the information transfer matrix. The row sum of the information transfer matrix is the total information transfer by the node to the rest of the network. The distribution plot shows that most of the nodes are not influential as they are sending little or no information to the rest of the network. There are few nodes with maximum influence on the network. The information-based influence can also be used to cluster the network and identify influential node within each cluster. In Fig. 6.1c, we show the clusters for the Twitter follower network obtained using spectral clustering method (38). For the 200 node Twitter network there are eight clusters. The two of the clusters are large (marked with green and red) and then there

are six clusters consisting of a single node. Within each cluster, one can identify the most influential or the hub node. The most influential node within a cluster is the one which sends the maximum amount of information to the rest of the nodes within the cluster.

6.2   Information Transfer and Optimal Placement of Actuators

In this section, we present some preliminary results on optimal placement of actuators in a fluid flow vector field. The problem of optimal placement of actuators and sensors was studied in (39) and there the authors used operator theoretic ideas to find the optimal locations. A Finite-dimensional analysis was carried out using the set-oriented methods and an optimization problem was posed which solves the optimal location problem. In this section, we will show, through simulations, that the proposed definition of information transfer can also be used to obtain similar optimal locations of actuators in a fluid flow system.

Given a non-linear system $\dot{x} = f(x)$, the problem is to find the optimal location of actuators in the state space for efficient control of the dynamics. For example, in building systems applications, the objective is to determine optimal locations of actuators in the form of vents/ducts for the control of temperature. For the control of contaminants from an oil spill in oceanographic flows, the goal is to determine the optimal location for the release of dispersant. The distribution of temperature in the room or the dispersant on the ocean surface is assumed to be modeled by a passive scalar density. The passive scalar is advected by the fluid flow velocity field. Under some simplifying assumptions of negligible buoyancy and diffusion, the evolution of the passive scalar density is modeled using the linear advection partial differential equation (PDE). In fact, the problem of control of temperature in a building system or prediction of contaminants in fluid flow is challenging because of the nonequilibrium nature of dynamics that are involved. The fluid flow velocity fields involved in these applications are quite complex with dynamics consisting of multiple equilibrium points, periodic orbits, limit cycles and chaotic attractors (40; 41). Although the point-wise evolution of a particle under the influence of the velocity field is complex, the evolution of density, modeling the ensemble behavior, is linear and is described by the linear advection partial differential equation (PDE).

There is an extensive amount of literature on actuator and sensor placement for a linear PDE with

(42) providing an excellent review of the results. Most of the methods for the optimization of sensors and actuators location involve finite dimensional approximation of the infinite dimensional system. In (43), sensor and actuator placement for diffusive and heat type partial differential equations are discussed. The placement problem for flexible structures is studied in (44). In (45), the provided analytical expression for the finite time controllability and observability gramain for the advection PDE. Selection criteria for the optimal location of actuators and sensors were proposed based on the maximization of gramians. In (39), we provided analytic expression for the infinite time controllability and observability for the advection PDE and proposed linear programming based approach to find the optimal locations of actuators and sensors.

We assume that the scalar density is advected by a vector field described by following differential equation:

$$\dot{x} = f(x) \tag{6.1}$$

where $x \in X \subset \mathbb{R}^{\mathcal{N}}$ is the finite dimensional spatial variable. The space X is assumed to be compact. Typically $\mathcal{N} = 2, 3$ in all our proposed applications. We assume that the vector field $f(x)$ satisfy following assumption to guarantee existence of solution on the compact space X.

Assumption 28. *We assume that the vector field generated by f has compact support in X and hence is complete in X. This guarantees the global in-time solution of the differential equation (6.1) in X (46).*

The evolution of the scalar density, $\rho(x, t)$, is governed by the following linear advection partial differential equation.

$$\frac{\partial \rho(x, t)}{\partial t} = -\nabla \cdot (f(x)\rho(x, t)), \tag{6.2}$$

where, $\rho(x, t) \in L^2(X)$ is the the infinite dimensional state belonging to state space $L^2(X)$. Let $A_k \subset X$ and $S_\ell \subset X$ be the locations of actuators and sensors respectively for $k = 1, \ldots, p$ and $\ell = 1, \ldots, q$. We make following assumption on the actuator and sensor sets A_k and S_ℓ respectively. The controlled evolution of linear advection PDE (6.2) with spatially located actuators and sensors

can be described as follows:

$$
\begin{aligned}
\frac{\partial \rho(x,t)}{\partial t} &= -\nabla \cdot (f(x)\rho(x,t)) + \sum_{k=1}^{p} \chi_{A_k}(x)u_k(x,t) \\
y_\ell(x,t) &= \chi_{S_\ell}(x)\rho(x,t), \quad k=1,\ldots,p, \quad \ell=1,\ldots,q
\end{aligned}
\tag{6.3}
$$

where $\chi_{A_k}(x)$ and $\chi_{S_\ell}(x)$ are the indicator functions for the set A_k and S_ℓ respectively. $u_k(x,t)$ is the control input for the kth actuator and y_ℓ is the output of the ℓth sensor. The control formulation of the advection PDE in Eq. (6.3) can be motivated as follows. Consider the application of oil spill control, where the control, in the form of dispersant, is released from multiple locations, say $A_k \subset X$, in the ocean. If the fluid flow field is modeled by the vector field $\dot{x} = f(x)$, then the evolution of the dispersant, $\rho(x,t)$, is modeled by Eq. (6.3). Similarly, the controlled evolution of temperature density in building system application will be modeled using Eq. (6.3). For building systems applications, the control input will consist of hot/cold air released from various vent locations in the room. Similarly for applications involving sensor placement, it is assumed that the sensor can make the measurement of the density $\rho(x,t)$ in the small region S_ℓ of the physical space.

In (39), we provided analytical expressions for the infinite time controllability and observability and proposed a linear program to solve the optimal placement problem. To arrive at the optimization problem, we discretized the state space to obtain a finite dimensional approximation of the Perron-Frobenius operator and used it to propose an optimization problem which solves the optimal placement problem. In order to obtain a finite-dimensional approximation of the continuous P-F operator, one considers a finite partition of the state space X as follows :

$$
\mathcal{X} = \{D_1, D_2, \ldots, D_L\}
\tag{6.4}
$$

such that $X = \cup_{i=1}^{L} D_i$ and $D_i \cap D_j = \phi$ for all $i \neq j$. With this partition, the finite dimensional approximation of the P-F semi-group is a $L \times L$ matrix P such that

$$
[P]_{ij} = \frac{m(f^{-1}(D_j) \cap D_i)}{m(D_i)}
\tag{6.5}
$$

where m is the Lebesgue measure. The ij^{th} entry of the finite dimensional P matrix gives the probability of transition from cell D_i to cell D_j. Hence $[P]_{ij} = \text{Prob}(z_{t+1} \in D_j | z_t \in D_i)$. Moreover, if $\mu_z = \begin{pmatrix} \mu_{z_1} & \mu_{z_2} & \cdots & \mu_{z_L} \end{pmatrix}$ is the density of z at time t, then

$$
\mu_{z'} = \mu_z P
\tag{6.6}
$$

is the density of z at time step $(t+1)$.

So far, in this thesis, we have defined the information transfer between the states of a dynamical system. To address the problem of optimal location of actuators, we need to define the information transfer between the cells D_i and study how much information one cell transfers to the other.

Let $\mu^t = \begin{pmatrix} \mu_1^t & \mu_2^t & \cdots & \mu_L^t \end{pmatrix}$ be the distribution on the state space at any time t. Then the entropy of the distribution is given by

$$H(\mu^t) = -\sum_{i=1}^{l} \mu_i^t \log \mu_i^t \tag{6.7}$$

We define the entropy of cell D_i at time t as $H(D_i) = -\mu_i^t \log \mu_i^t$. As the system evolves from time step t to time step $t+1$, the new distribution at time step $t+1$ is $\mu^{t+1} = \mu^t P$. Hence,

$$\mu_j^{t+1} = \sum_{i=1}^{L} \mu_i^t [P]_{ij}. \tag{6.8}$$

So the entropy transfer from cell D_i to cell D_j is $[T_{D_i \to D_j}]_t^{t+1} = -\mu_i^t [P]_{ij} \log(\mu_i^t [P]_{ij})$. Similarly, using higher powers of P-F matrix, one can define the information transfer over multiple time steps. For example, the information transfer from cell D_i to D_j, as the system evolves from time step t to step $t+2$ is $[T_{D_i \to D_j}]_t^{t+2} = -\mu_i^t [P^2]_{ij} \log(\mu_i^t [P^2]_{ij})$, where $[P^2]_{ij}$ is the ij^{th} entry of P^2. So the total information transfer over two time steps is $[T_{D_i \to D_j}]_t^{t+1} + [T_{D_i \to D_j}]_t^{t+1}$.

Now, the problem is to find the placement of minimum number of actuators so that one can control the entire state space. In terms of information transfer, one needs to identify the minimum number of cells such that there is non-zero information transfer from these cells to all the other cells. Let $e \in \mathbb{R}^{1 \times L}$ be a vector of length L with entries zero and one. Then one can write the optimal placement problem as

$$\min \quad ||e||_0$$
$$\text{subject to} \quad eT > 0$$
$$e_i \in \{0, 1\}$$

where T is the information transfer matrix such that $[T]_{ij}$ is the information transfer from cell D_i to

cell D_j. The optimization problem is non-convex but can be convexified via the following relaxation

$$\min \quad ||e||_1$$

$$\text{subject to} \quad eT > 0$$

$$e_i \in [0, 1]$$
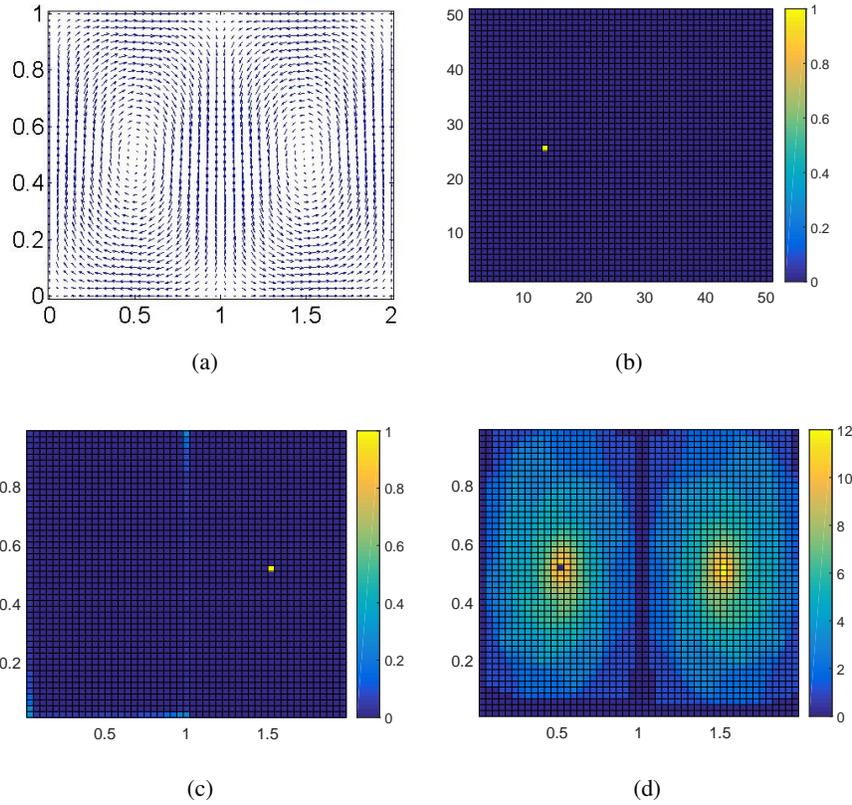
Figure 6.2   (a) Fluid flow vector field. (b)-(d) Eigenvectors corresponding to unit eigenvalue of the P-F matrix.

With this framework, we look at the problem of optimal location of actuators in a fluid flow vector field. The Double Gyre flow field is given by the following equation

$$\dot{x} = -\pi \sin(\pi x) \cos(\pi y), \quad \dot{y} = \pi \sin(\pi y) \cos(\pi x)$$

where $(x, y) \in X = [0, 2] \times [0, 1]$. The Double Gyre flow field is used as a toy model for oceano-graphic flow field (47). The vector field of the double-gyre fluid flow is shown in fig. 6.2(a). We discretize the state space into 50×50 cells and the corresponding finite-dimensional P-F matrix has
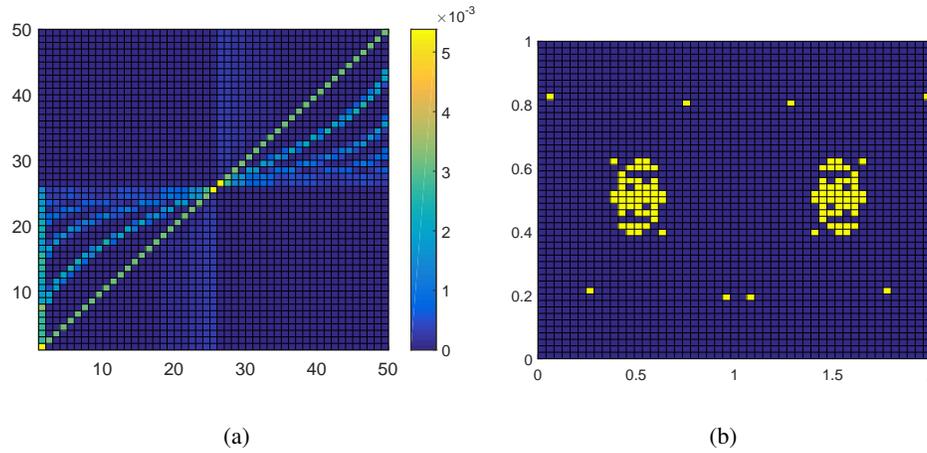
Figure 6.3   (a) Information transfer between the cells. (b) Optimal location of 200 actuators.

three eigenvalues equal to one. The corresponding eigenvectors give the invariant sets in the state space and this is shown in fig. 6.2(b)-(d).

In fig 6.2(a) we show the information transfer between the first 50 cells and in fig. 6.2(b) the location of 200 actuators is shown. We find that most of the actuators are located around the center and this is consistent with the results obtained in (39). One difference is that in this simulation, we use the total information transferred over 500 time steps to find the optimal location, whereas in (39) we had used an infinite series of the powers of P-F matrix to compute the optimal locations. In future, we hope to have a more rigourous connection between the information transfer framework and the optimal placement problem.

CHAPTER 7.   INFLUENCE CHARACTERIZATION IN POWER SYSTEMS

7.1   Introduction

The power system is one of the largest man-made systems and has played a pivotal role in the development of human society in the past century. Control and maintenance of power system is a challenging task, where different components are connected in a complicated network and one of the primary challenges is to maintain the stability of power system. Being a large network with complicated topology, it is very difficult to identify sources of instability, hence an *a priori* knowledge of the sources of instability will allow one to take appropriate control actions at those sources and thus maintain stability. Again, there are many types of instability associated with the power system, namely, voltage instability, rotor angle instability and frequency instability (48). It is of great importance to know the type and cause of instability which is leading to system collapse. In particular, if one can identify the generator or states of the generator which is at the root of instability, then proper control and planning action can be taken in time to prevent any failures in the system.

Power system stability problem can be addressed at two different time scales, as long term and short term stability. For long term stability, load increase, change in network topology, slow acting equipment such as tap changing transformer, generator current limiter are considered; whereas, short term stability deals with transient phenomena such as an occurrence of a fault, stalling of an induction motor and rotor angle instability. Long term stability analysis is mostly model based and interested readers can look in (49; 50; 51) and the references therein. Study of short term voltage stability is still in nascent phase. Various dynamic models for power system have been developed for time domain analysis of the transient behavior of the system (52; 53; 54). Recent development also incorporates model free approach, where data based analysis of short term stability are studied (55).

Here, we work on the problem of stability characterization in power system. In particular, we use

the newly developed notion of information transfer to characterize which generators and also which dynamic state(s) of the generators are contributing most towards system instability and oscillations. We analyze the IEEE 9 bus network to show that the information transfer from generators and states of the generators to the most unstable mode (or critical mode) and most complex mode identifies the sources of instability. We also model the load and identify the generator which has maximum influence on the load.

7.2   Stability Issues in Power Systems

Power system stability is an area of interest both for academic research and industrial innovation. As defined in (48), *power system stability is the ability of a system to regain operating equilibrium after being subjected to any physical disturbance*. In a primitive setup, power systems were operated with conservative margins so as to keep the system well within stability limits. Advancement of measurement and protection technologies has enabled stressed operation of the power system and the modern grid is operated close to voltage instability. Thus we are able to economically transfer more power with existing infrastructure. In such a setup, it has become more crucial to understand the nature and cause of instability and the behavior of overall dynamics of the system. Various models, theories and simulation tools have been developed to understand the nature of instabilities. These simulation tools enable us to capture a variety of scenarios and the behavior of the system, which helps in characterizing the system at a macro level. However, detailed study of system dynamics and participation of different components of a power system in overall stability is not included in these studies.

For stability studies, it is essential to understand both the type of instabilities arising in the system and also how different machines/component and their dynamics are contributing to those instabilities.

For instance, at the macro level, a voltage instability may arise in one area due to rotor angle instability in another area. Hence, it is crucial, to understand the evolution of system and instabilities for timely control and planning actions. In (56), the authors study the mechanism of voltage instability, where they discuss how different areas and components contribute towards instability and the study was based on the concept of participation factor (28; 57).

In conventional power system operation, any change in load or any physical disturbance trigger change in generator operating point. Thus it is essential to take into account generator dynamics and their contribution leading to system instability. In this study, we explore dynamic behavior of the system at different operating conditions. The objective is to understand the nature of instabilities that can arise in the system and then characterize the contribution of different components towards it. In our study, we further analyze the contribution of specific generator states, namely exciter voltage, rotor angle etc. to particular instabilities evolving in the system. This study will help us to develop a theoretical framework for mapping different instabilities to the different physical parameters of the system.

7.3   Modelling of a Power Network

A power system is a complicated network with different components, with the most important being the generators, transformers, transmission lines, substations, distribution lines, distribution transformers and loads. All these components make a power system a highly nonlinear system and thus a complicated system to analyze from the theoretical point of view. So, a common approach to the analysis of a power system is to linearize the system along its trajectories. Here we will consider dynamics of a generator and its local control, along with load dynamics considered as aggregate induction motor load.

7.3.1   Generator Model

The model used in this work is based on the modeling described in (58). The power network is described by a set of differential algebraic equations (DAE) and the power system dynamics is divided into three parts: differential equation model describing the generator and load dynamics, algebraic equations at the stator of the generator and algebraic equations describing the network power flow. We consider a power system model with n_g generator buses and n_l load buses. The generator dynamics at each generator bus can be represented as a 4th order dynamical model:

$$\frac{d\delta_i}{dt} = \omega_i - \omega_s$$

$$\frac{d\omega_i}{dt} = \frac{T_{m_i}}{M_i} - \frac{E'_{q_i}I_{q_i}}{M_i} - \frac{(X_{q_i} - X'_{d_i})}{M_i}I_{d_i}I_{q_i} - \frac{D_i(\omega_i - \omega_s)}{M_i}$$

$$\frac{dE'_{q_i}}{dt} = -\frac{E'_{q_i}}{T'_{do_i}} - \frac{(X_{d_i} - X'_{d_i})}{T'_{do_i}}I_{d_i} + \frac{E_{fd_i}}{T'_{do_i}}$$

$$\frac{dE_{fd_i}}{dt} = -\frac{E_{fd_i}}{T_{A_i}} + \frac{K_{A_i}}{T_{A_i}}(V_{ref_i} - V_i)$$

$$(7.1)$$

The algebraic equations at the stator of the generator are:

$$V_i\sin(\delta_i - \theta_i) + R_{s_i}I_{d_i} - X_{q_i}I_{q_i} = 0$$

$$E'_{q_i} - V_i\cos(\delta_i - \theta_i) - R_{s_i}I_{q_i} - X'_{d_i}I_{d_i} = 0$$

$$\text{for} \quad i = 1, \ldots, n_g.$$

$$(7.2)$$

The network equations corresponding to the real and reactive power at generator and load buses are shown below.

$$I_{d_i}V_i\sin(\delta_i - \theta_i) + I_{q_i}V_i\cos(\delta_i - \theta_i) + P_{L_i}(V_i)$$

$$- \sum_{k=1}^{\bar{n}} V_iV_kY_{ik}\cos(\theta_i - \theta_k - \alpha_{ik}) = 0$$

$$I_{d_i}V_i\cos(\delta_i - \theta_i) - I_{q_i}V_i\sin(\delta_i - \theta_i) + Q_{L_i}(V_i)$$

$$- \sum_{k=1}^{\bar{n}} V_iV_kY_{ik}\sin(\theta_i - \theta_k - \alpha_{ik}) = 0$$

$$\text{for} \quad i = 1, \ldots, n_g.$$

$$P_{L_i}(V_i) - \sum_{k=1}^{\bar{n}} V_iV_kY_{ik}\cos(\theta_i - \theta_k - \alpha_{ik}) = 0$$

$$Q_{L_i}(V_i) - \sum_{k=1}^{\bar{n}} V_iV_kY_{ik}\sin(\theta_i - \theta_k - \alpha_{ik}) = 0$$

$$\text{for} \quad i = n_g + 1, \ldots, n_g + n_l.$$

$$(7.3)$$

Here δ_i, ω_i, E_{q_i}, and E_{fd_i} are the dynamic states of the generator and correspond to the generator rotor angle, the angular velocity of the rotor, the quadrature-axis induced emf and the emf of fast

acting exciter connected to the generator respectively. The algebraic states I_{d_i} and I_{q_i} are the direct-axis and quadrature-axis currents induced in the generator respectively. Each bus voltage and its angle are denoted by V_i and θ_i. The parameters $T_{m_i}, V_{ref_i}, \omega_s, M_i,$ and D_i are the mechanical input and machine parameters applied to the generator shaft, reference voltage, rated synchronous speed, generator inertia, and internal damping. The stator internal resistance is denoted by R_{s_i} and $X_{q_i}, X_{d_i},$ X'_{d_i} are the quadrature-axis salient reactance, direct-axis salient reactance and direct-axis transient reactance. The exciter gain and time-constant are given by K_{A_i} and T_{A_i}.

A power system stabilizer (PSS), that acts as a local controller to the generator is designed based on the linearized DAEs. The input to the PSS controller is $\omega_i(t)$ and PSS output, $V_{ref_i}(t)$, is fed to the fast acting exciter of the generator. An IEEE Type-I PSS is considered here which consists of a wash-out filter and two phase-lead filters. The transfer function of PSS is given by

$$\frac{\Delta V_{ref_i}(s)}{\Delta \omega_i(s)} = k_{pss}\frac{(1 + sT_{num})^2}{(1 + sT_{den})^2}\frac{sT_w}{1 + sT_w} \tag{7.4}$$

where k_{pss} is the PSS gain, T_w is the time constant of wash-out filter and T_{num}, T_{den} are time constants of phase-lead filter with $T_{num} > T_{den}$.

Elimination of the algebraic variables by Kron reduction generates a reduced order dynamic model given by $\Delta \dot{x}_g = A_{gg}\Delta x_g + E_1 \Delta \tilde{u}$ where $\Delta x_g \in \mathbb{R}^{7n_g}$ and $\Delta \tilde{u} \in \mathbb{R}^{n_g}$.

7.3.2 Load Model

In this subsection, we discuss the load model for the power network and derive a dynamic linear model with both the generator and load dynamics. For adding dynamics at load buses, we consider an aggregate induction motor load representing distributed smaller loads (59). Third order dynamics are considered for representing induction motor, with quadrature axis Thevenin voltage (e'_{q_i}), direct axis Thevenin voltage (e'_{d_i}) and angular speed of rotor (ω_i) being the dynamic state variables. With these, the dynamic equations are

$$T'_{0_i}\frac{de'_{q_i}}{dt} = -T'_{0_i}(\omega_i - \omega_{0_i})e'_{d_i} - e'_{q_i} - (L_{s_i} - L'_i)i_{d_i}$$

$$T'_{0_i}\frac{de'_{d_i}}{dt} = -T'_{0_i}(\omega_i - \omega_{0_i})e'_{q_i} - e'_{d_i} + (L_{s_i} - L'_i)i_{q_i} \tag{7.5}$$

$$\omega_i\left(\frac{2H_i}{\omega_{0_i}}\right)\frac{d\omega_i}{dt} = (e'_{q_i}i_{q_i} + e'_{d_i}i_{d_i}) - T_{\text{nom}_i}\left(\frac{\omega_i}{\omega_{0_i}}\right)^D_i$$

$$\text{for} \quad i = n_g + 1, \ldots, n_g + n_l$$

where T'_{0_i} is the transient rotor time constant, T_{nom_i} is the nominal value of mechanical torque, ω_{0_i} is the synchronous speed of the induction motor, L_{s_i} is the synchronous inductance, L'_i is the motor transient inductance, H_i is the inertia constant and D_i is the load model exponent.

The algebraic equations for the induction motor are

$$e'_{d_i} - je'_{q_i} + (R_{a_i} + jL'_i)(i_{d_i} + ji_{q_i})$$

$$= V_i\cos\theta_i + jV_i\sin\theta_i\text{for} \quad i = n_g + 1, \ldots, n_g + n_l \tag{7.6}$$

where R_{a_i} is the stator winding resistance. From equation (7.6), i_{q_i} and i_{d_i} can be expressed in terms of V_i and θ_i and substituting these expressions of i_{q_i} and i_{d_i} in (7.5), the dynamic equations can be written in the standard linear system form. Again, the load voltages can be expressed in terms of the generator voltages (58) and using the network flow equations (7.3) one obtains the coupled linear model of the generator and load as

$$\begin{pmatrix}\Delta\dot{x}_g \\ \Delta\dot{x}_l\end{pmatrix} = \begin{pmatrix}A_{gg} & 0 \\ A_{lg} & A_{ll}\end{pmatrix}\begin{pmatrix}\Delta x_g \\ \Delta x_l\end{pmatrix} + \begin{pmatrix}E_1 \\ 0\end{pmatrix}\Delta\tilde{u} \tag{7.7}$$

where $\Delta x_l \in \mathbb{R}^{3n_l}$.

7.4 Influence Characterization in IEEE 9 Bus System

As discussed earlier, identification of sources of instability in power network is an important problem because knowledge of the sources of instability would allow implementation of correct corrective measures at the proper nodes of the power network and thus enable better control of power system.

In this section, we describe our approach to identify the states and the generators which are most responsible for instability in a power network. The idea is based on the newly developed notion of information transfer(60), (61). In particular, we use the notion of information transfer to define influence in a dynamical system and analyze how each state (or generator) influences the instability of a power system.

In this work, we analyze how individual states (or generators) in a power network influence the stability of the power network.
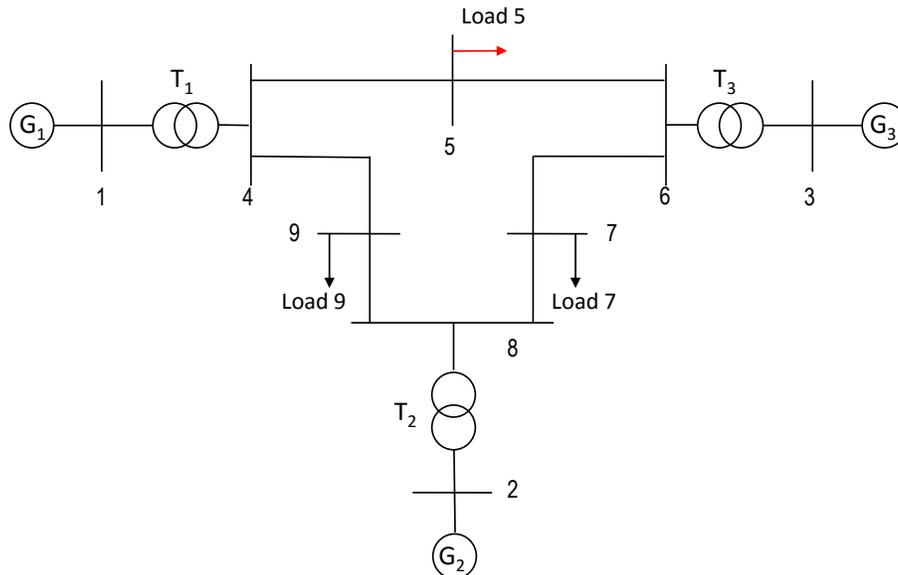
Figure 7.1   IEEE 9 bus network

In particular, we identify the generator or set of generators which are most likely to cause instability and oscillations in the IEEE 9 bus system with dynamic load at bus 5 (fig. 7.1), and also identify the generator or set of generators, where one needs to take action to account for load changes. This allows us to identify which generator is *participating* most in causing instability and which generator can cause the most changes on the load side. As of now, some, and not all, of these questions were answered in terms of participation factors (28) (57). The concept of participation factor is based on the computation of left and right eigenvectors and is not mathematically rigorous. Moreover, it cannot

account for the influence of one state on another state in a system. But our concept of information transfer is much more rigorous (60; 22) and is applicable to a wide variety of applications. In particular, it can not only quantify the influence of states of a system on a particular mode but also can characterize how each state (or subspace) can influence any other state (or subspace). Another advantage of the information transfer measure over participation factor is the fact that this can be applied to nonlinear systems as well. Moreover, the developed notion of information transfer is also related to classical concepts of structural controllability and observability (60) and these connections can be used to locate the nodes where actuation or sensing can be performed.

In this work, we compare the results obtained using our developed framework of information transfer with the existing concept of participation factor and demonstrate what new information our approach provides.

We use a seventh order model for each of the three generators and a third order model for the load. For load dynamics, we considered induction motor load at bus 5 with constant power factor. This results in a 24th order continuous time linear model for the power network. For computation of information transfer, we discretized the continuous time system with discretization step $\delta t = 1$ second.
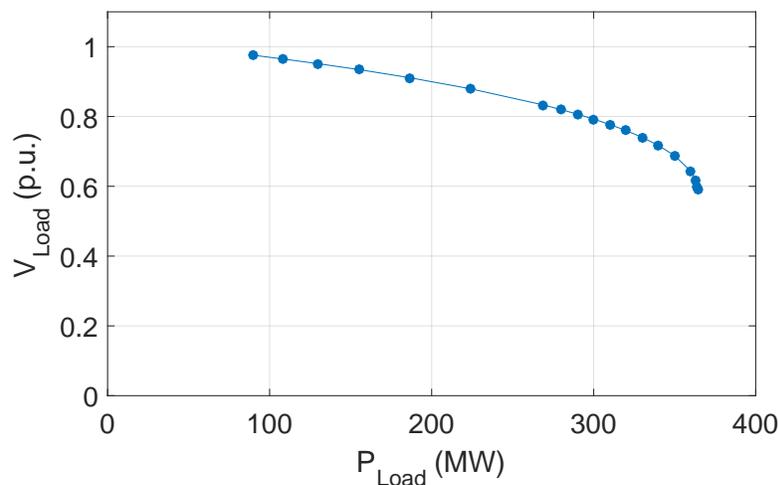
Figure 7.2   P-V curve with operating points

In order to study the behavior of the system in different operating conditions, power flow solution was obtained at different operating points with variation in load (MW) at bus 5, as represented in Fig. 7.2. Load at bus 5 is assumed to be an induction motor load with constant power factor. G_2 and G_3 contribute towards this increase in load proportional to their machine inertia and G_1 connected at slack bus supply the remaining power to balance the power flow. As a base case, the system operates with $P_{load} = 90$MW at bus 5. As we increase the loading, the system finally collapses after $P_{load} = 364.1$MW. Here voltage collapse is observed due to stress on the power system. 19 different operating points are considered (Fig. 7.2) for different loading conditions between the base case and voltage collapse point and at each operating point a steady state solution is obtained using Newton-Raphson power flow solution. At each operating point, these solutions are used to obtain the 24th order continuous time linear model.

7.4.1 Contribution to most unstable mode

We study the contribution of each generator and individual states of each generator to instability at different operating points along the PV curve.
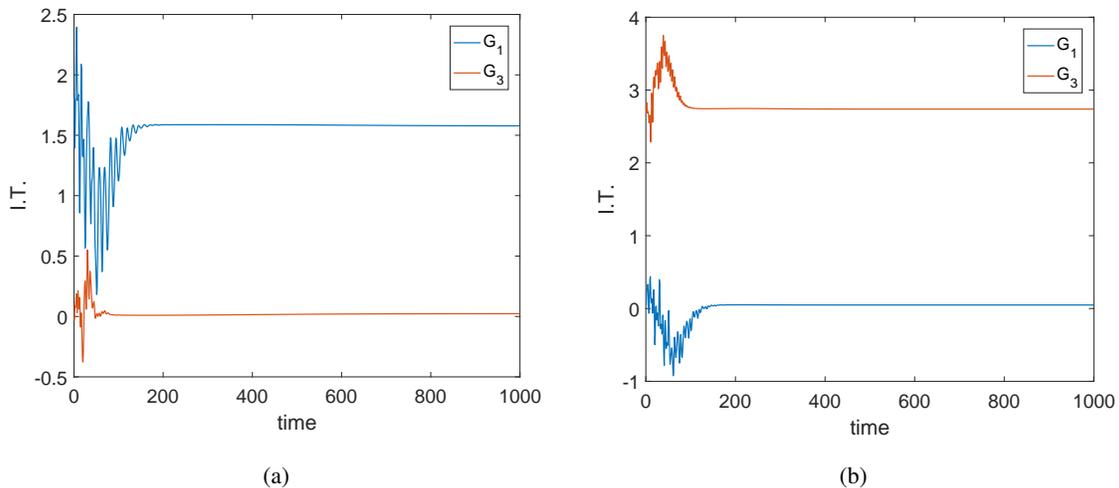
Figure 7.3   Information transfer from generators 1 and 3 to most unstable mode for (a) $P_{load} = 90$ MW. (b) $P_{load} = 364.1$ MW.

The most unstable mode is the eigenvector corresponding to the eigenvalue closest to the imaginary axis for the continuous time system and for a discrete time system, it corresponds to the eigen-

value closest to the unit circle. We look at the information transfers from generators 1 and 3 to this most unstable mode. In Fig. 7.3(a) we show the information transfer from the two generators to the most unstable mode when the load is $P_{load} = 90$ MW. We find that G_1 is transferring more information to the most unstable mode and hence conclude that at this operating point, G_1 is contributing most to the instability of the system. This is in agreement with the participation factor, as indicated in Table 7.1. In Fig. 7.3(b) the information transfer from the same generators to the most unstable mode is shown for the case when $P_{load} = 364.1$ MW. It is observed that near the nose point, that is, voltage collapse, the contribution of G_3 towards instability is more than the contribution of G_1.

Table 7.1   Participation to most unstable mode

Gen.\Op. Pt.	1	19
G_1	0.606	0.0014
G_3	0.394	0.9984

This is again, in accordance with the participation factor analysis, as depicted in Table 7.1. We do not plot the information transfer from G_2 as the information transfer and participation factor for G_2 remains negligibly small compared to G_1 and G_3 throughout the PV curve. This is a valid outcome as the system is getting stressed at bus 5 due to increase in load which is in turn fed mostly by G_1 and G_3 as shown in Fig. 7.1 in the network topology. From Fig. 7.3(a) and (b), we find that near the nose point G_3 contributes more towards instability when compared to G_1, whereas, near initial operating points, contribution of G_1 is more. So there is some operating point in between, where G_3 becomes more influential than G_1.

In Fig. 7.4(a) we plot the steady state information transfer from G_1 and G_3 to the most unstable mode, as we move along the PV curve. We find that in the initial stages G_1 is contributing more towards instability than G_3. However, as we move along the PV curve towards the nose point, G_1's contribution decreases and when the load is around 270 MW, the influence of G_3 becomes more than that of G_1. In Fig. 7.4(b), we plot the participation factors of G_1 and G_3 to the most unstable mode. Here also, initially G_1 is participating more than G_3 and from around 270 MW G_3 starts participating more than G_1. So the switch of influence is also indicated correctly by our information transfer measure.
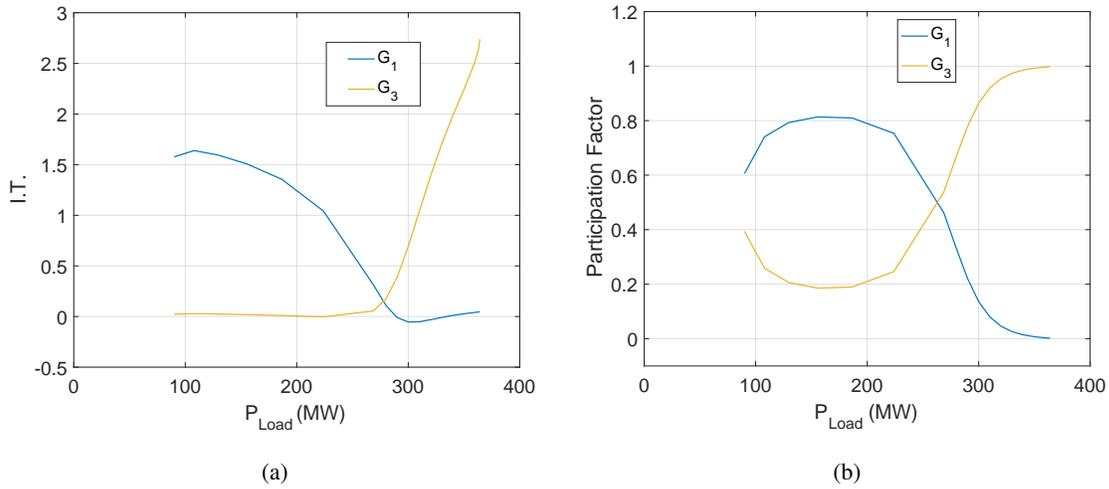
Figure 7.4   (a) Steady state information transfer to most unstable mode from G_1 and G_3 along operating points. (b) Participation factors of G_1 and G_3 to most unstable mode along operating points.

From the above simulations, we found that in the initial loading condition on the PV curve, G_1 contributes more to the voltage instability than G_3. So next, we look at individual states of G_1 and identify which state of G_1 is most responsible for instability. Fig. 7.5(a) shows the information transfer from the four states of generator 1 to the most unstable mode when $P_{load} = 90$ MW.

We find that the emf of the fast acting exciter (E_{fd}) is most responsible for instability. This is physically consistent because due to the increase in load(MW), the system is approaching voltage collapse (voltage instability) and in a synchronous generator, exciter emf is responsible for mainte-nance of bus voltage in operating range. Similarly, for the operating point where $P_{load} = 364.1$ MW, we again find that E_{fd} of G_3 contributes most to instability. This is shown in Fig. 7.5(b).

Fig. 7.6 shows the steady state information transfer from the exciter emf of G_1 and G_3 to the most unstable mode. We find that in the initial part of the PV curve, E_{fd} of G_1 participates more to instability than E_{fd} of G_3 and from around $P_{load} = 270$ MW, the E_{fd} of G_3 contributes more to instability than that of G_1. This is in exact accordance with the fact that from around $P_{load} = 270$ MW, G_3 starts contributing more towards voltage instability than G_1.
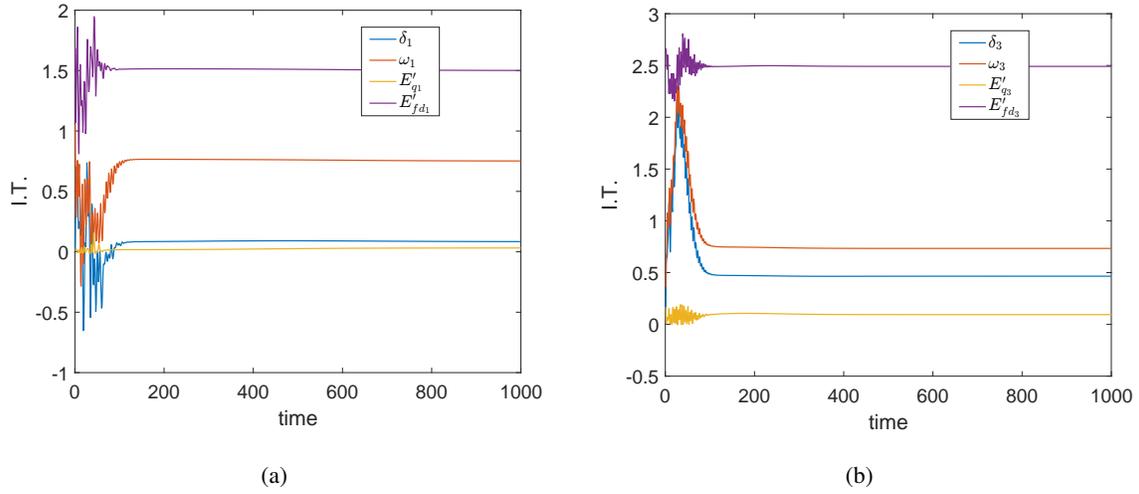
Figure 7.5 (a) Information transfer from individual states of G_1 to most unstable mode for $P_{load} = 90$ MW. (b) Information transfer from individual states of G_3 to most unstable mode when $P_{load} = 364.1$ MW.
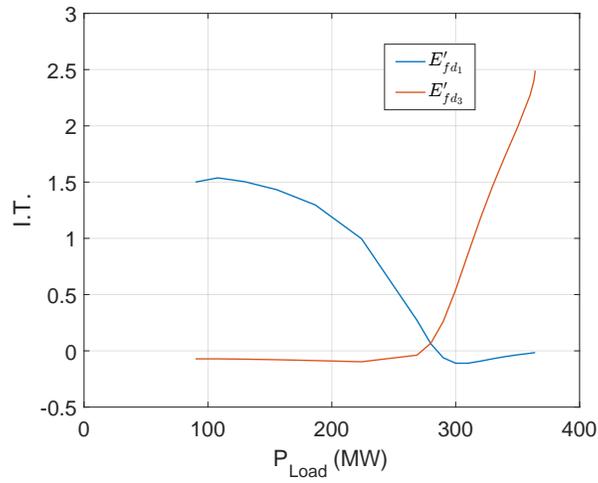
Figure 7.6 Steady state information transfer from E_{fd} of G_1 and G_3 to most unstable mode along operating points.

7.4.2 Information transfer to the complex modes

From linear system theory, we know that the eigenvalues with non-zero imaginary parts are responsible for oscillations of a linear system. In this subsection, we identify the states of the generators in the power network which are responsible for oscillations in the network. We only present the simulation results for generator 3 at operating points 1 and 19 and similar conclusions hold for other generators and operating points. We consider the *most complex mode* (eigenvalue with largest imaginary part) of the system and look at how the states of the generators influence this mode.
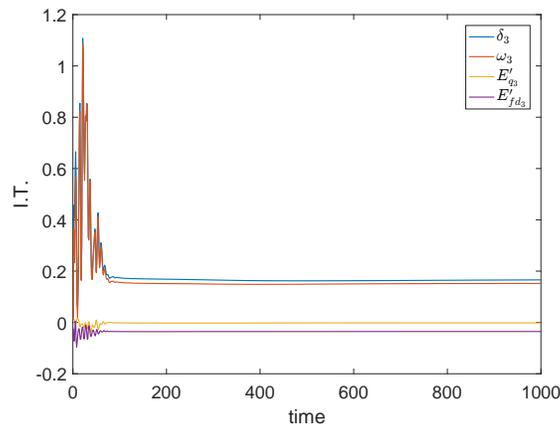
(a)
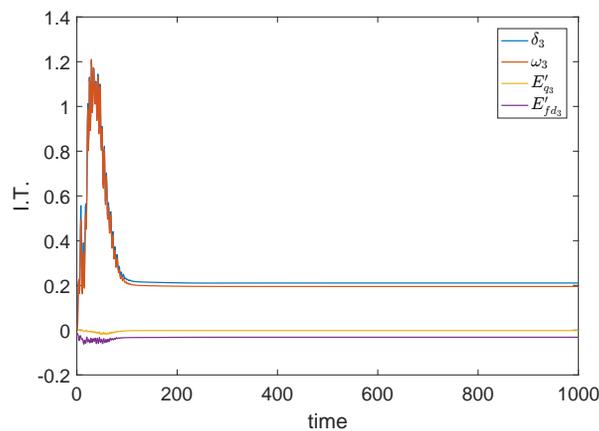
(b)

Figure 7.7 Information transfer from states of G_3 to complex mode (a) for operating point 1. (b) for operating point 19.

Fig. 7.7(a) shows the information transfer from the individual states of generator 3 to the most

complex mode, for operating point 1. The discretized eigenvalue corresponding to this mode is $\lambda = 0.8010 + 0.5003i$. From the information transfer graph, it can be seen that the most information is transferred by those states which correspond to the rotor angle(δ) and rotor angular speed(ω).

Table 7.2   Participation to most complex mode

States\Op. Pt.	1	19
δ_3	0.1678	0.2107
ω_3	0.1524	0.1951
E_q'	−0.0018	−0.0017
E_{fd}'	−0.0353	−0.0315

This is in accordance with the physical behavior and intuitions of a power system, as these are the states which are responsible for the angular stability of the system and thus lead to short term rotor angle instability. Similarly, for operating point 19, from fig. 7.7(b), we find that the same variables are transferring most information to the most complex mode. In this case the most complex mode corresponds to an eigenvalue of $\lambda = 0.5947 + 0.7060i$. In fact, the trend that the rotor angle and the rotor angular speed is transferring most information to the most complex mode is followed through all the operating points on the PV curve. This information is validated with participation factor analysis, which is shown in Table 7.2. We find that at both operating points participation of generator state δ and ω is higher as compared to other two states and this is revealed by our measure as well.

7.4.3   Information flow from generators to load

Participation factor gives the participation of a state to any mode, but it does not give any information about how a state affects another state. Our information transfer measure is more general in this aspect and it gives a quantitative information about how a state affects another state. In this subsection, we investigate how the individual generators contribute to the load dynamics. The load is modeled as an induction motor and has its own dynamics. We consider a third order model for the load with the states being the quadrature axis Thevenin voltage, direct axis Thevenin voltage and rotor angular speed. The load is connected to the generators through the power network and in this present set of simulations, we considered the dynamic load to be at bus 5. To analyze how each generator

influences the load dynamics, we consider the information transfer from the individual generators to the load.

(a)   (b)
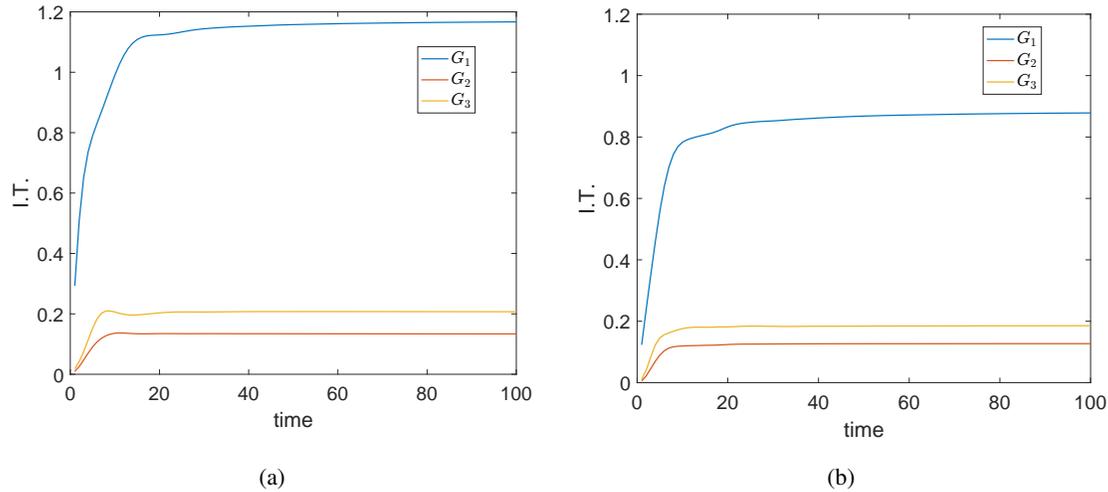
Figure 7.8   Information transfer from the generators to the load (a) at operating point 1. (b) at operating point 19.

Fig. 7.8(a) and (b) shows the information transfer from the generators 1, 2 and 3 to the load at operating points 1 and 19. It can be seen that in both the cases, it is generator 1 that transfers the most amount of information to the load. This is due to the fact that G_1 is close to load at bus 5 and in the current modeling G_1 is considered as a slack bus. So we conclude that of the three generators, it is generator 1 that has the most influence on the load dynamics. In fact, throughout the PV curve, we found that generator one has the most influence on the load. But from the steady state values of the information transfer, it can be seen that as one approaches the nose point (operating point 19, signifying instability of the power network), the information transfer, and hence influence, of the generator 1 decreases.

In fig. 7.9, we plot the steady state values of the information transfer from generator 1 to the load at all the 19 operating points. It can be clearly seen that the information transfer decreases as the load is increased. One thing to be noticed is that the information transfer curve almost exactly mimics the PV curve and even shows the dip that the PV curve has near the nose point. As to why the steady state information transfer from G_1 to the load mimics the PV curve is left for future investigations.
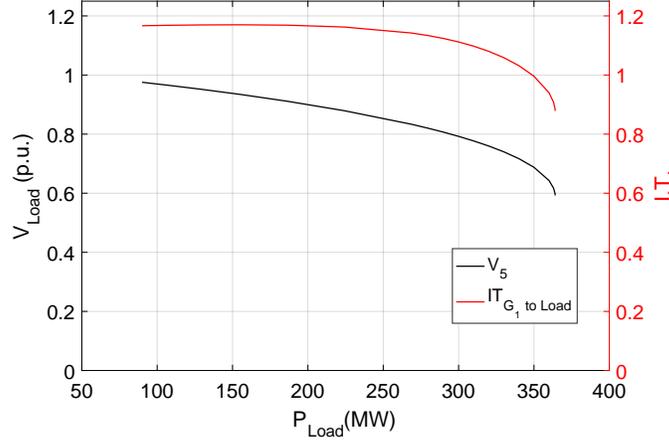
Figure 7.9   Steady state values of information transfer from G_1 to load along the operating points (red curve) and PV curve (black curve). Here V_5 is the voltage at bus 5.

7.5   Stability Characterization from State to State Information Transfer

In the previous example, stability analysis of power network was based on the information transfer from state to mode. However, identification of the state variables using information transfer to the modes or participation factor needs the computation of eigenvalues and eigenvectors which may be computationally expensive. In this section, we provide a procedure of identifying the states responsible for instability from state to state information transfer.

Consider a linear system

$$x(t+1) = 0.7x(t) + y(t) + \xi_x(t)$$
$$y(t+1) = \mu y(t) + \xi_y(t)$$

(7.8)

where $\mu \in [.1, .99]$ and $\xi_x(t)$ and $\xi_y(t)$ are independent and identical Gaussian noise of unit variance. The eigenvalues of the system are $(0.7, \mu)$ and hence as μ approaches 1, the system approaches instability. The instability occurs due to y dynamics and as μ increases, the entropy of y increases rapidly. Hence, the information transfer from y to x also increases rapidly. This is shown in Fig. 7.10.   Conversely, if the information transfer from some state (subspace) to any other state (subspace) increases rapidly, it can be concluded that the system is approaching instability. The point to be noted is that the rapid increase in the information transfer happens when the system is still operating in the stable
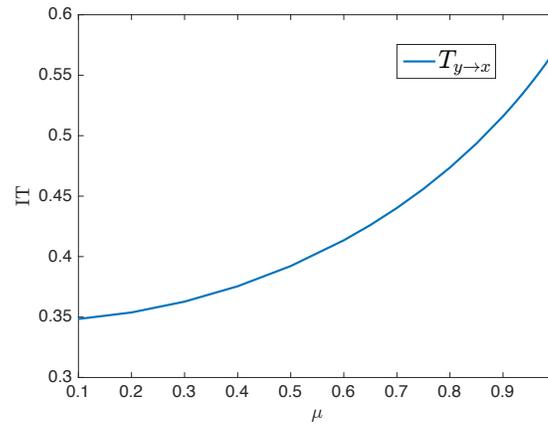
Figure 7.10   Information transfer increases rapidly as the system approaches instability.

zone. Hence, one can use information transfer to predict the onset of instability and hence one can take preventive measures before the system becomes unstable.

7.5.1   3 Bus system

Consider a power network with two generators and a load, as shown in figure 7.11(a). The load is modelled as an induction motor in parallel with a constant PQ load. The system is modelled as a four dimensional dynamical system with the state being generator angle (δ_g), generator angular velocity (ω), the load angle (δ_l) and magnitude of load voltage (V).
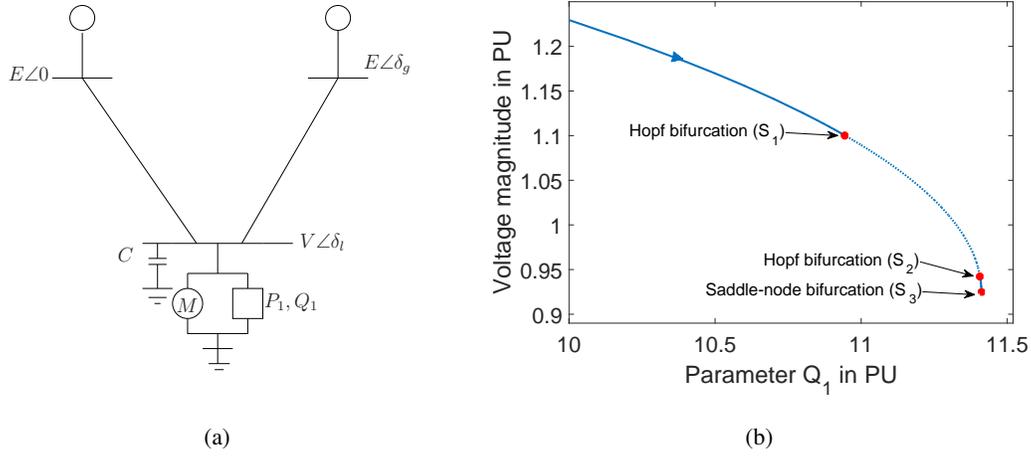
Figure 7.11 (a) 3-bus system. (b) Critical points of the system

The dynamic equations for the system are

$$\dot{\delta_g} = \omega \tag{7.9}$$

$$\dot{\omega} = 16.66667\sin(\delta_l - \delta_g + 0.08727)V$$
$$-0.16667\omega + 1.88074 \tag{7.10}$$

$$\dot{\delta_l} = 496.87181V^2 - 166.66667\cos(\delta_l - \delta_g$$
$$-0.08727)V - 666.66667\cos(\delta_l - 0.20944)V$$
$$93.33333V + 33.33333Q_1 + 43.33333 \tag{7.11}$$

$$\dot{V} = -78.76384V^2 + 26.21722\cos(\delta_l - \delta_g$$
$$-0.01241)V + 104.86887\cos(\delta_l - 0.13458)V$$
$$+14.52288V - 5.22876Q_1 - 7.03268 \tag{7.12}$$

For detailed analysis of the system equations we refer the interested reader to (62; 63).

The above power network has three *critical points*, namely S_1, S_2 and S_3, as shown in figure 7.11(b). At S_1 and S_2, a pair of imaginary eigenvalues cross the imaginary axis and at S_3, a real eigenvalue becomes zero. Hence, the system becomes unstable at S_1, remains unstable from S_1 to S_2, then regains stability after S_2 and again becomes unstable at S_3. It is known that the instability at S_1 is angle instability and the instability at S_3 is voltage instability.

7.5.2 Information Transfer and Participation Factor

Participation matrix P (64) for a linear system

$$\dot{x} = Ax \tag{7.13}$$

is defined as

$$P = [p]_{ki} = \mathbf{u}_k^i \mathbf{w}_k^i \tag{7.14}$$

where $\mathbf{u}_k^i$ ($\mathbf{w}_k^i$) is the kth component of the ith left (right) eigenvector corresponding to eigenvalue λ_i of the power system A matrix (Eq. 7.13). These eigenvector are assumed to be normalized i.e.,

$$\mathbf{u}_i^\top \mathbf{w}_j = 1, \ \text{ if } \ i = j; \ 0 \ \text{ otherwise.}$$

Participation factor p_{ki} measure the relative participation of kth state variable in the ith mode. Hence, larger the participation factor of a state to a particular mode, larger the contribution of that state to that mode.

For calculating the participation factors of each state of the 4-bus system (7.9)-(7.12), we look at the linearized dynamics at the operating point $Q = 1141.1$ MVar. The rationale behind choosing this operating point is the fact that if Q is increased further, the system undergoes Saddle Node Bifurcation (voltage instability). It is well known (63) that at SNB, it is the load voltage V that participates the most in the most unstable mode and this is shown in table 7.3.

Table 7.3 Participation and Information Transfer to most unstable mode

State.\Index	Participation Factor	Information Transfer
δ_g	0.0137785	0.04
ω	0.00046	0.0094
δ_l	0.00072	-0.0189
V	0.9850	1.39884

Information transfer from the states to the most unstable mode, at this operating point, is shown in table 7.3. For calculating the information transfer we discretized the linearized continuous-time model with $\Delta t = 0.001$ and calculated the steady state information transfer from the states to the most unstable mode. It is observed that both information transfer and participation factor identify

the load voltage as the state responsible for instability. Not only does information transfer identify the load voltage as the state most responsible for instability, but it can be seen that the order of participation/influence is same with both participation factor and information transfer. So far, we see that participation factor and information transfer are doing the same thing. But one of the big advantages of the information transfer measure is the fact that one can look at information transfer from any combination of states to any combination of states or from any combination of state to any output. Hence, in a system with more that two generators, one can identify which generator is responsible for instability. For more details on the relationship between participation factor and information transfer, we refer the interested reader to (65).

7.5.3 Stability and state to state information transfer

It is known that in the 3 bus network, there are both angle and voltage instability. In particular, the instability that occurs at S_1 (Fig. 7.11) is angle instability and the instability that occurs at S_3 is voltage instability (63).
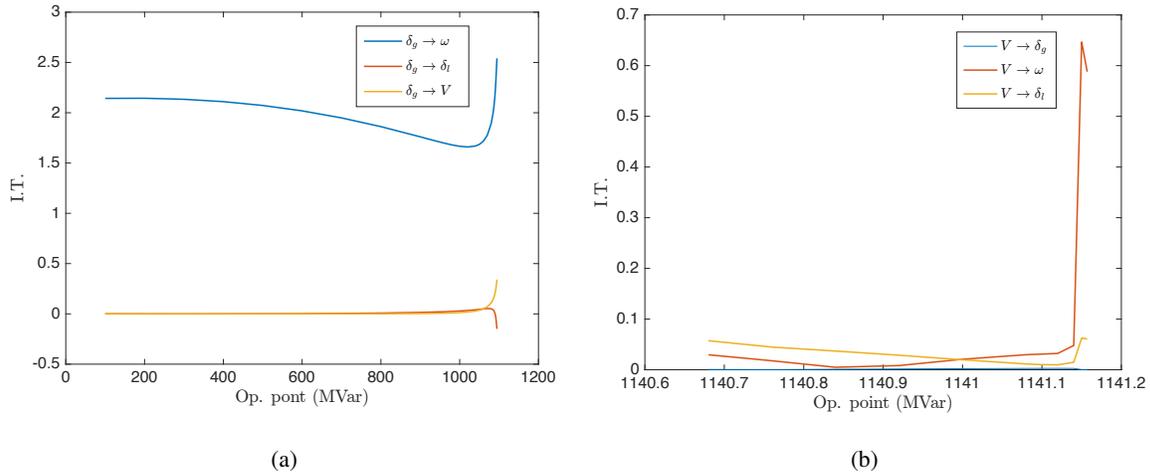
Figure 7.12 (a) IT from δ_g before S_1. (b) IT from V before S_3.

In Fig. 7.12(a) we show the information transfer from the angle variable of the generator to all the other states, as the system approaches the first Hopf bifurcation. It can be seen that the information transfer from the angle variable shows a sudden increase. This is consistent with the fact that at S_1, the

system undergoes voltage instability and it is the angle variable that causes the instability. Similarly, as the system approaches voltage instability (S_3), the information transfer from the angle variable (V) to all the other states shows the sudden increase. This is shown in Fig. 7.12(b). Hence one can identify the states which are responsible for the instability of the power network. Participation factor can also identify the same, but participation factor has some drawbacks (66) and moreover, it requires the computation of eigenvalues and eigenvectors. But the information transfer framework does not require the computation of eigenvalues and eigenvectors. Moreover, as discussed later, one can do similar computation from time series data as well, which is a huge advantage of the information transfer framework.

7.6   IEEE 39 Bus System

In this section, we consider the IEEE 39 bus system, the line diagram of which is shown in Fig. 7.13. The model used is based on the modeling described by (7.1).
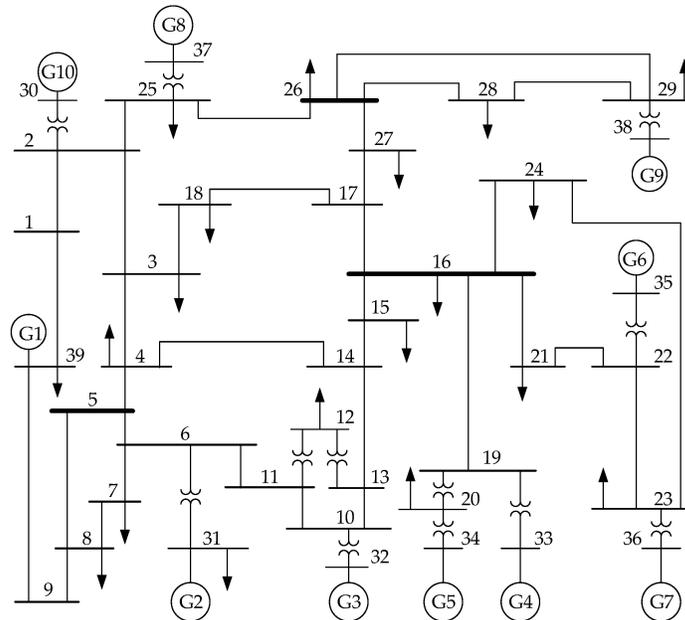
Figure 7.13   IEEE 39 bus system.

The objective is to use the information transfer measure to identify the generator and the states of the generator which are responsible for instability. For identifying the states responsible for instability,

the concept of participation factor is generally used. But one drawback of participation factor is the fact that it can not compute the participation of a combination of states to any particular mode (except in some special cases). However, our formulation of information transfer is free from such issues and one can compute the information transfer from any combination of states to any other combination of states. As such, one can combine the states of each generator together and look at the information transfer between the generators.

In the first set of simulations, we look at the information transfer between the generators and identify the generator the information transfer from which shows a sudden increase, as the network approaches instability. For simulation purpose, we linearized the system along the PV curve and used the linearized model to compute the information transfer between the generators. The discretization step used was 0.2 secs.

Simulation results identified generator 10 as the generator which is responsible for instability. This is because, the information transfer from generator 10 showed a sudden rise, as the network approached instability. Information transfer from generator 10 to generators 1 and 2 is shown in Fig. 7.14(a) and from the figure it can be clearly seen that the information transfer from generator 10 starts to increases rapidly. For comparison, we also plot the information transfer from generators 5, 8 and 9 to generators 3, 4 and 6 respectively in Fig. 7.14(b). It can be clearly seen in Fig. 7.14(b) that the information transfers between these generators does not change that much, thereby implying that these generators are not responsible for instability.

Once we have identified the generator which is most responsible for instability, we zoom in to the generator and identify the states of the generator which are responsible for instability. We consider only the four states of the generator and not the states of the PSS. In particular, we study the information transfer from the states of generator 10 to other generators and identify the states of generator 10 the information transfer from which to other generators show the sudden increase.

In Fig. 7.15(a) we plot the information transfer from the individual states of generator 10 to generator 2. From the figure, it can be seen that the information transfer from the angle and the angular speed variables show the sudden increase, whereas the other transfers remain almost same throughout the PV curve. Hence, we conclude that it is the angle and the angular speed variables
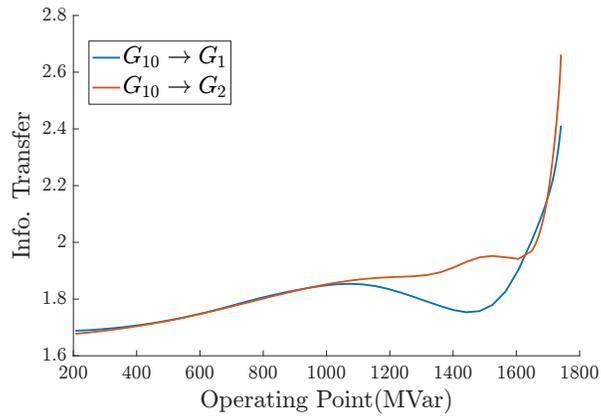
(a)
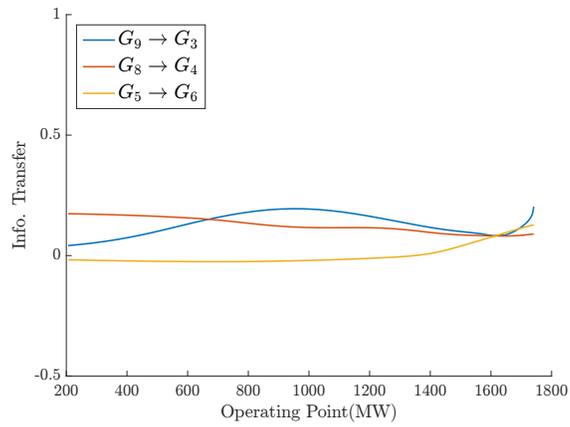
(b)

Figure 7.14 (a) Information transfer from generator 10 to generators 1 and 2. (b)Information transfer from generators 5, 8 and 9 to generators 3, 4 and 6 respectively.
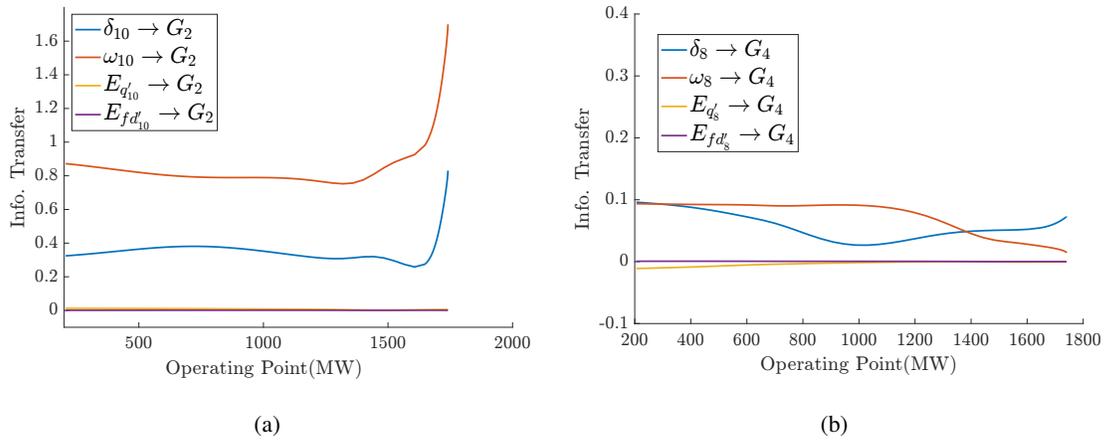
Figure 7.15   (a) Information transfer from generator 10 to generators 1 and 2. (b)Information transfer from generators 1 and 2 to generator 10.

which are most responsible for instability. This fact is also verified by participation factor analysis, where it was found that the participation factor of the angle and angular speed of generator 10 to the most unstable mode is 0.47 each. Hence, these variables are almost 94% responsible for instability. For comparison, we also plot the information transfer from the states of generator 8 to generator 4 in Fig. 7.15(b). As expected, we find that the information transfer does not change much throughout the PV curve, and hence reaffirming the fact thath generator 8 is not responsible for instability. The same conclusion holds for the states of the other generators as well.

CHAPTER 8.   INFORMATION TRANSFER AND PHASE TRANSITION IN DYNAMICAL SYSTEMS

Phase transition is most commonly used to describe passage from one state like solid to another state like liquid. However, over time, phase transition has been viewed as an order-disorder transition. In this context, phase transitions are ubiquitous in nature and have been found in behaviour of bacteria (67), locusts (68) and also in financial market (69). Phase transition is also quite common in complex systems featuring ensembles of dynamically interacting components.

Characterization of phase transition has been a problem of interest and one of the major breakthroughs was achieved by Landau (70), where he connected phase transition with breaking of symmetry. The importance of symmetry breaking was first realized by Pierre Curie when he concluded *asymmetry is what creates a phenomenon*. He also discussed what is now known as *spontaneous symmetry breaking* (71). There are other types of symmetry breaking as well, like explicit symmetry breaking (72).

Another important question, related to phase transition is the following : Is it possible to predict if a system is moving towards phase transition? In (73) the authors used mutual information as a measure to identify phase transitions in complex dynamical systems. They found that during phase transition in Vicsek model (67), the mutual information reaches a peak. More recently, in 2013, in (74), the authors showed that in a kinetic Ising model transfer entropy (26) reaches a peak in the disordered phase, just before the disorder-order transition. However, in a dynamical system setting, it has been shown in (60) that transfer entropy may not give the correct causal structure. In particular, it has been shown that in a dynamical system there may be a non-zero transfer entropy from a state x to a state y, but in reality x dynamics does not affect the y dynamics, neither directly nor indirectly. Hence, while studying phase transitions in a dynamical system setting, it is logical to use a measure

which gives the correct causal structure and influence characterization.

In this work, we analyze phase transitions in dynamical systems which are associated with symmetry breaking, with this newly developed concept of information transfer (60; 61). In particular, we use the information transfer between the states (subspaces) of the underlying dynamical system and show, via examples, that the information transfer can act as an indicator for symmetry breaking phase transitions. In fact, we show that the information transfer can also predict the phase transition in advance. In particular, we analyze phase transition in the standard double-well potential setting and use the intuitions obtained to analyze switching of lobes in Duffing oscillator and Lorenz system. We further show how information transfer can be used to identify phase transition in complex networks, where phase transition can occur due to the formation of Giant Components or condensation of edges (75).

8.1 Information Transfer and Phase Transition

The abrupt change in the behavior of a system, as a consequence of a continuous change in a given parameter, is well-known in literature (76). In particular, as the given parameter is changed, for some critical value of the parameter trajectories makes a sudden transition from one type of behavior to a different type. This is known as phase transition. Another characterization of phase transition, in terms of symmetry breaking, is due to Landau (70). He proposed that phase transition is accompanied by breaking of *symmetry*. For example, consider a particle in a symmetric double-well potential, as shown in Fig. 8.1.

When the particle is located at the local maxima at 0, it is in a state of complete symmetry. However, any small disturbance will cause the particle to enter into one well or the other and the initial symmetry is broken. This is a classic example of symmetry breaking and phase transition.

Consider another classic example where ice melts to form water. We consider a very simplistic "two-dimensional" ice crystal. We assume there are six molecules of H_2O in the ice crystal, which are arranged in a perfect hexagon as shown in Fig. 8.2. We consider the symmetry group of the ice crystal.

From basic geometry, the hexagonal crystal is symmetric w.r.t reflection about the line ℓ and also
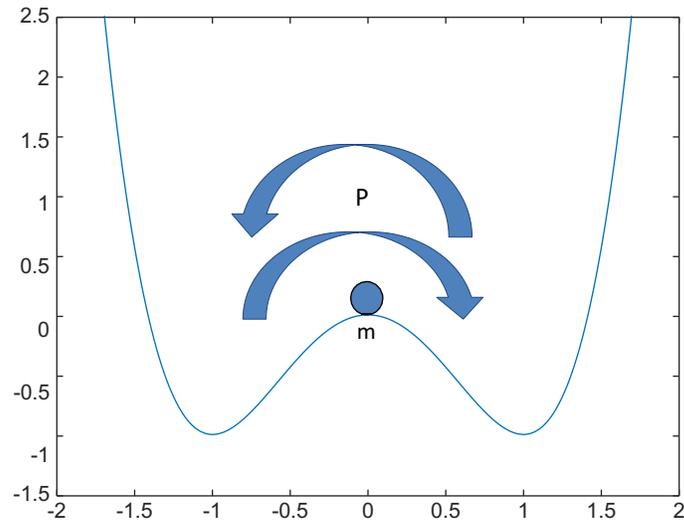
Figure 8.1   Particle in a symmetric double-well potential.
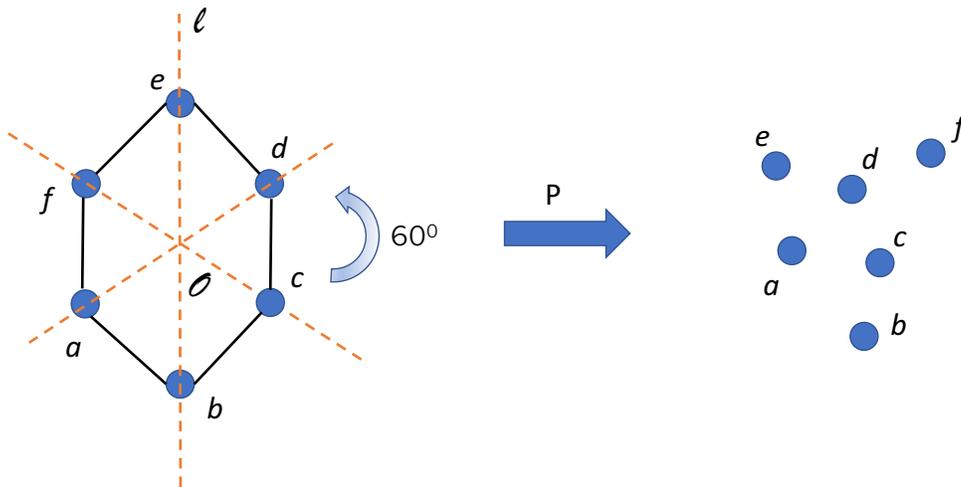
Figure 8.2   Phase transformation from ice to water.

w.r.t. rotations by $60°$ about $\mathcal{O}$. Let r denote the reflection about ℓ and s denote anti-clockwise rotation by $60°$ about $\mathcal{O}$. Then it r and s satisfy the following relations.

$$r^2 = s^6 = e \tag{8.1}$$

where e is the identity operation. In addition to (8.1), to specify the symmetry group completely, one needs to identify the relation between r and s. For this, consider

$$rsr^{-1}(a) = rs(c) = r(d) = f = s^{-1}(a) \tag{8.2}$$

Hence, the symmetry group (G) for the ice lattice can be written as a group generated by r and s such that

$$G = \langle r, s | r^2 = s^6 = e, rsr^{-1} = s^{-1} \rangle \tag{8.3}$$

The reflection about the other lines can be generated from combination of r and s. This is in fact, the dihedral group D_6 and can be written as a semi-direct product (77) of $\mathbb{Z}_2$ and $\mathbb{Z}_6$, that is, $D_6 = \mathbb{Z}_6 \rtimes_\phi \mathbb{Z}_2$, where $\mathbb{Z}_6$ represent rotation, $\mathbb{Z}_2$ represent reflection and ϕ can be thought of as the *relation* that connects rotation and reflection. In particular, ϕ is a group homomorphism from $\mathbb{Z}_2 \rightarrow \text{Aut}(\mathbb{Z}_6)$ where $\text{Aut}(\mathbb{Z}_6)$ is the automorphism group of $\mathbb{Z}_6$ and the action is defined by conjugation. That is,

$$\phi : \mathbb{Z}_2 \rightarrow \text{Aut}(\mathbb{Z}_6)$$

$$h \mapsto \phi_h$$

$$\phi_h(n) = hnh^{-1}$$

Now, as the ice melts into water, the rotational symmetry is completely broken and this breaking of symmetry corresponds to the phase transition. In fact, water has complete rotational and translational symmetry (hence infinite symmetry).

In the context of complex networks, phase transition can occur due to the emergence of giant connected components and/or condensation of edges etc. (75). The emergence of giant components or condensation of edges also lead to a breaking of initial symmetry and thus this kind of phase transition can also be explained in terms of the symmetry groups of the network and the breaking of its symmetry.

In this work, we will consider phase transition which results from this breaking of symmetry and show how information transfer between the states (subspaces) of a dynamical system can be used to identify phase transitions.

8.2   Coupled Oscillators in Double Well Potential

Consider a dynamical system of N coupled oscillators whose equations of motion are given by

$$m_k \ddot{\theta}_k = -\frac{\partial V}{\partial \theta_k}(\theta_k) - \mathcal{L}_k \theta - d\dot{\theta}_k, \quad k = 1, \cdots, N \tag{8.4}$$

where m_k, θ_k and d_k are the mass, angular position and damping co-efficient of the kth oscillator, V is the potential and $\mathcal{L}_k$ is the kth row of the Laplacian matrix $\mathcal{L}$. The potential V is assumed to be a double well, as shown in Fig. 8.3(a). In particular, we choose $V(\theta) = E(C\theta^4 - \theta^2)$, with $E = 2$ and $C = 0.5$.

The phase portrait of a single oscillator, with E and C as specified before, is shown in Fig. 8.3(b).

We consider $N = 2$ and initialize one of the particles in the potential well at -1 and the other in the potential well at $+1$. Phase transition, in this case, refers to the phenomenon when one of the particles, due to the influence of the other particle, escapes from its well and enter the other well. We assume the particle placed in the potential well at -1 has mass $m_1 = 100$ and the other particle placed in the potential well at $+1$ has mass $m_2 = 10$. As such, under the influence of m_1, m_2 may escape from its well and enter the other well, thus resulting in phase transition. This is shown in Fig. 8.4.

Depending on the damping coefficient d, phase transition may occur once or multiple times. In all the subsequent discussion, we assume the Laplacian $\mathcal{L} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and consider discretized linear model at each point on the state space for computation of information transfer. The discretization we considered was Euler approximation with discretization step $\delta t = 0.05$. The points on the state space were generated by solving the differential equation (8.4), with step size 0.005. We also assume that the damping co-efficient $d_1 = d_2$ for all the simulations.

Case I: Only one phase transition

In the first set of simulations, we set $d = 5$ and with this damping, oscillator 2 (O_2) makes only one transition from its own potential well and settles down at the minimum of the second well. This
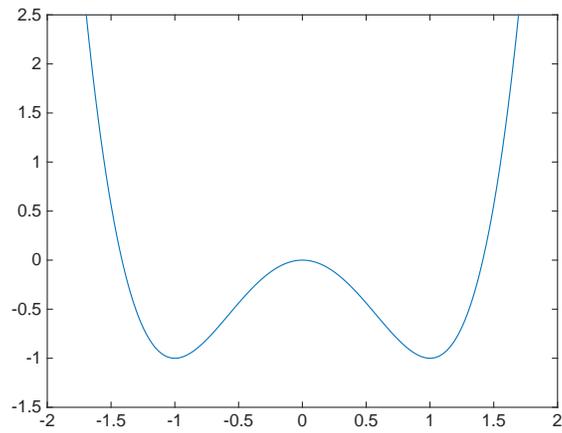
(a)
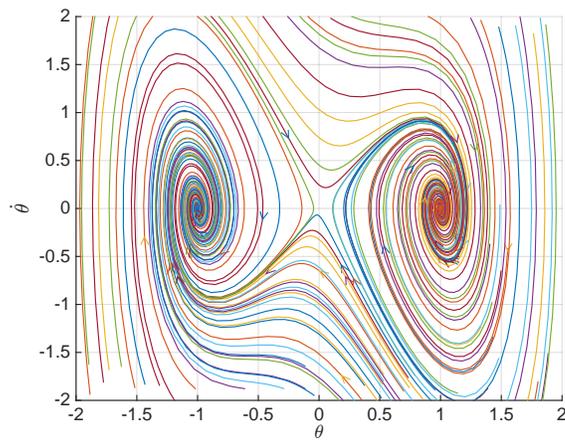
(b)

Figure 8.3   (a) Double well potential. (b) Phase portrait of a single oscillator.
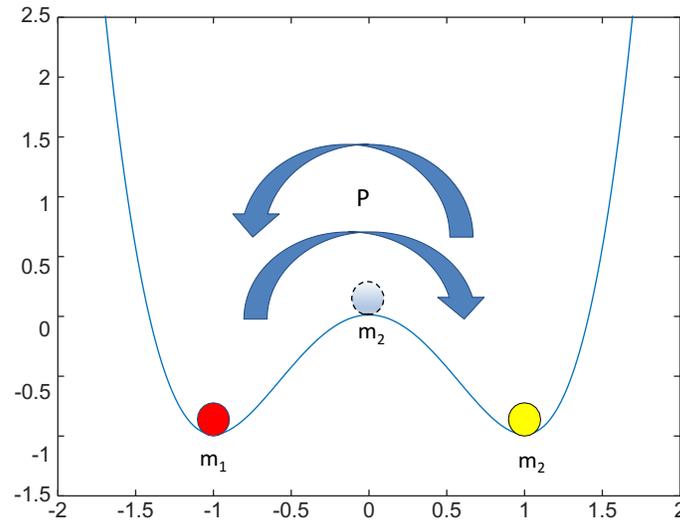
Figure 8.4   Oscillator 1 (marked by red), with mass $m_1 = 100$ remains in its well and oscillator 2 (marked by yellow) jumps from one well to the other, thus leading to phase transition.

is shown in Fig. 8.5.

The mass, and hence inertia, of oscillator 1 (O_1) is much larger compared to O_2 and as such, O_1 pulls O_2 out of the potential well at $+1$, thus resulting in phase transition. In the process, the energy of O_2 increases and at one point it gains enough energy to cross the barrier. Intuitively, this process can be thought of as O_2 being unstable, w.r.t. to the well at $+1$, and it enters the potential well at -1. As O_2 becomes more and more unstable, its own entropy increases rapidly and reaches a maximum near the local maxima of V at 0. As such, the information transfer to O_1 also increases rapidly and the information transfer peaks near the local maxima at 0. This is shown in Fig. 8.5 and the result obtained is similar to the findings in (74; 73). Hence, information transfer $T_{O_2 \to O_1}$ serve as an indicator for phase transition.

Case II : Multiple phase transitions

Based on the evidence of *Case I*, in cases, where oscillator 2 crosses the barrier multiple times, one should expect the information transfer to also peak multiple times. Not only that, the number of crossings should exactly match the number of peaks of the information transfer from O_2 to O_1.

To simulate multiple phase transitions, we decreased the damping coefficient d to $d = 0.05$ and simulated the system for sixty seconds. In these sixty seconds, O_2 crossed the barrier eight times and
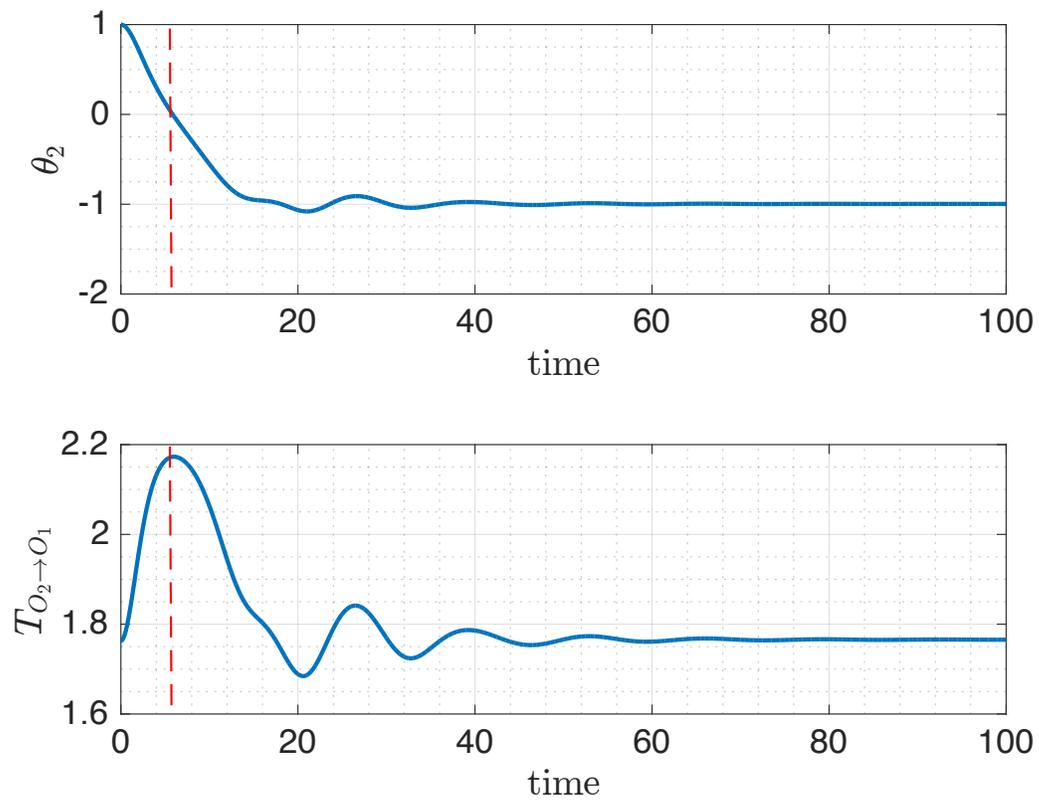
Figure 8.5 θ_2 and information transfer when there is only one phase transition.
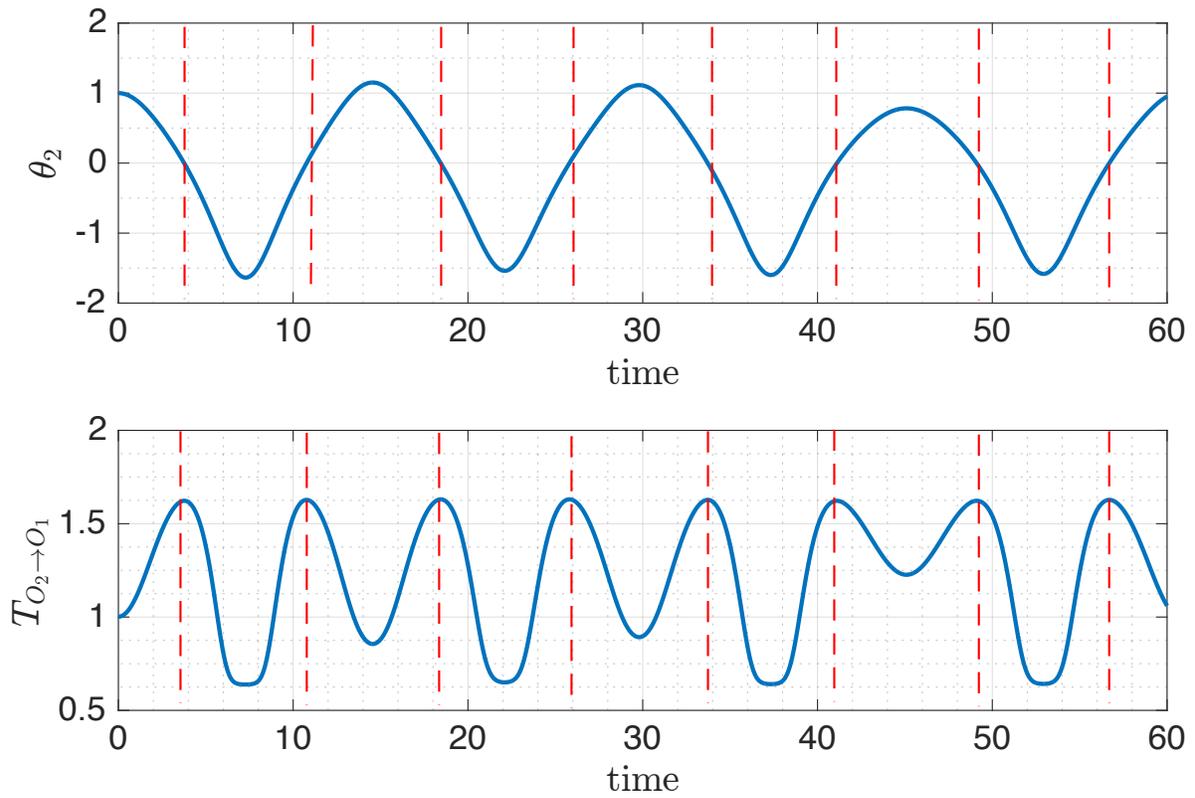
Figure 8.6 θ_2 and information transfer when there are multiple phase transitions.

the information transfer $T_{O_2 \rightarrow O_1}$ also peaked eight times and the peaks coincided with the crossing of the barrier of O_2. This is shown in Fig. 8.6. Hence, this establishes the fact that information transfer can indeed be used as an indicator for phase transition.

Case III : Almost a second transition

An interesting case is a situation when O_2 crosses over to the potential well at -1 and has enough energy to almost reach the global maxima at 0 for a second time, but does not have enough energy to cross over to the potential well at $+1$. In this case we can say that *a second transition almost occurred* (Fig. 8.7). It is evident that as O_2 nears the maxima for the second time (when it fails to cross over),
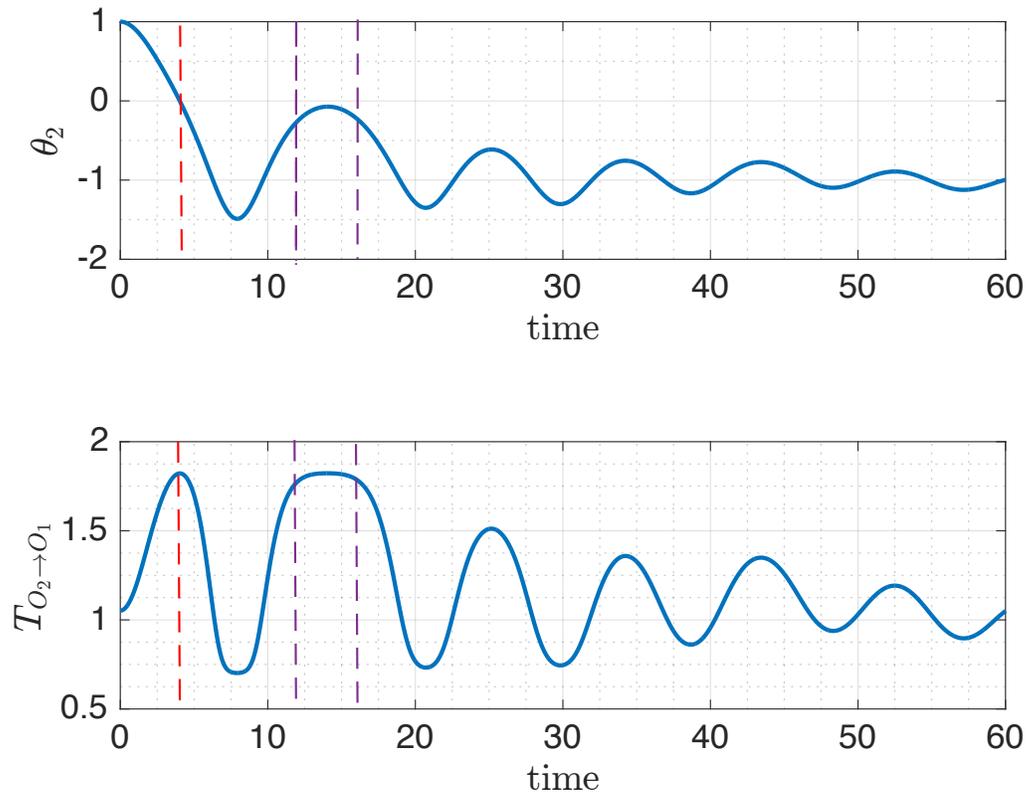
Figure 8.7 θ_2 and information transfer when there is only one phase transition and there is almost another transition.

its dynamics is very slow (that is, its velocity is small) and it remains near the maxima for a much longer time. As such, its entropy also remains high for a much longer time, as compared to the case when O_2 does cross the barrier. Hence, in this case, the information transfer should also peak, but

the difference lies in the fact that the information transfer from O_2 to O_1 is maintained at this peak value for a much longer time, and hence instead of a sharp peak, one should expect a plateau kind of a trajectory. In Fig. 8.7, we plot the information transfer $T_{O_2 \to O_1}$ and find that this is indeed the case and hence, in this case also information transfer spills out the information that there was almost a second phase transition.

8.3 Duffing Oscillator

A duffing oscillator is a second order nonlinear dynamical system which is used to model a class of damped and forced oscillations. The dynamic equation of double-well Duffing oscillator is

$$\ddot{x} + \delta\dot{x} + \alpha x + \beta x^3 = \gamma \cos(\omega t) \tag{8.5}$$

where $x(t)$ is the position variable and $\delta, \alpha, \beta, \gamma$ and ω are constants. Depending on these parameters, the Duffing oscillator can exhibit chaotic behaviour (78; 79). The equation (8.5) can be written in state space form as

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\psi} \end{pmatrix} = \begin{pmatrix} y \\ -\delta y - \alpha x - \beta x^3 + \gamma \cos\psi \\ \omega \end{pmatrix} \tag{8.6}$$

with $\psi(0) = 0$.

We considered $\delta = 0.2$, $\alpha = -1$, $\beta = 2$, $\gamma = 0.3$ and $\omega = 1$. With these values of the parameters, the double-well Duffing oscillator exhibits chaotic behavior as shown in Fig. 8.8. The behavior of the double-well Duffing oscillator is similar to the particle in a double-well potential, barring the chaotic nature of motion. As such, one can talk about phase transition in a Duffing oscillator in a way similar to the particle in a double-well potential and we say that there is a phase transition when the trajectory shifts from the right half plane (R.H.P) to the left half plane (the other way round) in the phase space.

For studying the phase transition in the Duffing oscillator, we look at the information transfer from x to y ($T_{x \to y}$) for the system given in (8.6). In particular, we track how the information transfer $T_{x \to y}$ changes, as the trajectory shifts from right half plane to the left half plane (L.H.P.) and otherwise. For computation of information transfer, we considered discretized linear model corresponding to (8.6) at
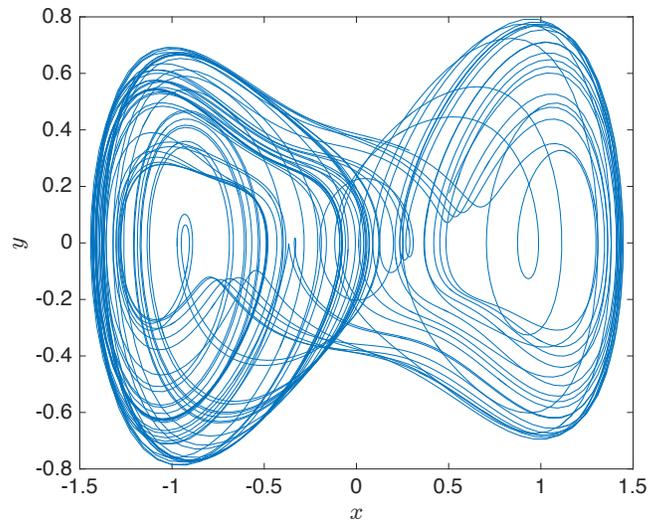
Figure 8.8   Phase portrait of Duffing oscillator
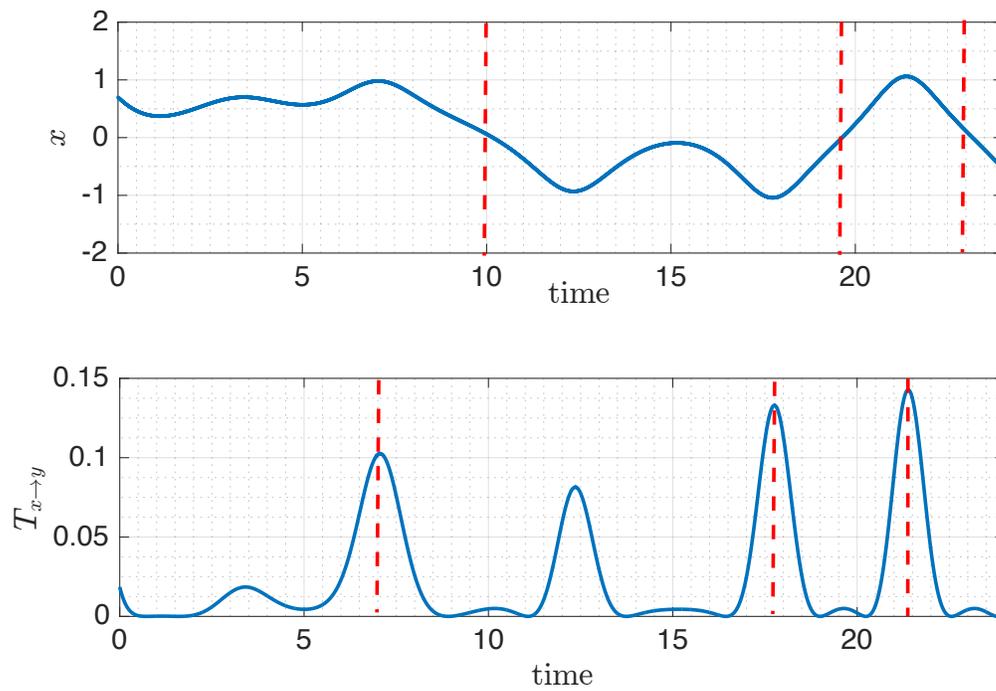
Figure 8.9   Information transfer and phase transition in Duffing oscillator.

each point on a trajectory. The discretization was Euler discretization, with $\delta t = 0.5$ and the trajectory was obtained by solving the continuous time system (8.6). As with the example of coupled oscillators in a double well potential, in the Duffing oscillator example too we found that the information transfer $T_{x \to y}$ peaks sharply before the trajectory shifts from R.H.P. to L.H.P. or from L.H.P. to R.H.P. This is shown in Fig. 8.9. A big difference from the coupled oscillator example which is to be noted is the fact that in the Duffing oscillator example, the information transfer peaks before the trajectory crosses over. The reason for this is left for future investigation. However, the takeaway point is that information transfer can be used to predict the crossing over and hence it can be used for prediction, though for small future times, of future of a chaotic system.

8.4 Lorenz Attractor

The Lorenz system was first proposed by Lorenz (80) in the context of weather prediction and it was only later that it was realized that the Lorenz system possesses such rich dynamical structures (81; 82; 83). The Lorenz system is a three dimensional nonlinear system with state space equations given by

$$
\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= xy - \beta z
\end{aligned}
\tag{8.7}
$$

For certain values of the parameters and initial conditions, the Lorenz system exhibits chaotic solutions and the Lorenz attractor is a set of chaotic solutions of (8.7). For $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$ the chaotic attractor is shown in Fig. 8.10.

The Lorenz attractor has two lobes, on either side of the plane $y = 0$ and similar to coupled oscillators, shifting of the trajectories from one lobe to the other can be thought of as a form of *phase transition*[1].

The system (8.7), with parameter values set to as described, was simulated from an initial condition so that it generates the attractor and information transfer from y to x ($T_{y \to x}$) was computed using

---

[1]Note that this is not phase transition in the truest sense, but analogous to coupled oscillators in a double well where the shifting of lobes is similar to transition of one oscillator from one well to the other.
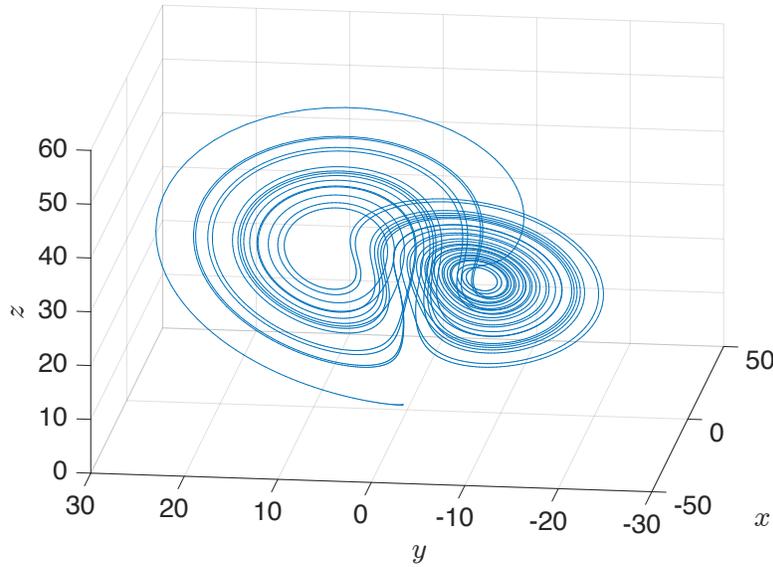
Figure 8.10   Phase portrait of Lorenz system.

discretized linear model along the trajectory.

In the Lorenz system, it was again observed that the information transfer $T_{y \to x}$ peaks just before the trajectory crosses the $y = 0$ plane. Hence, in this example also one can use information transfer to predict the transition from one lobe to another. This is consistent with the findings of (74), where the authors show that transfer entropy (26) peaks just before a phase transition in an Ising model. However, the authors have shown in (60) that in a dynamical system setting, measures of causality like transfer entropy, directed information (11) and Granger causality (8) fail to capture the true causal structure. Moreover, in (60; 61) the authors proposed a new definition of information transfer and showed that the new definition of information transfer captures the true notion of causality in a dynamical system. These simulations establish the fact that this new definition of information transfer (definition [6]), though inherently different from the other measures of causality, do behave similarly to them when it comes to prediction of transitions in complex dynamical systems.
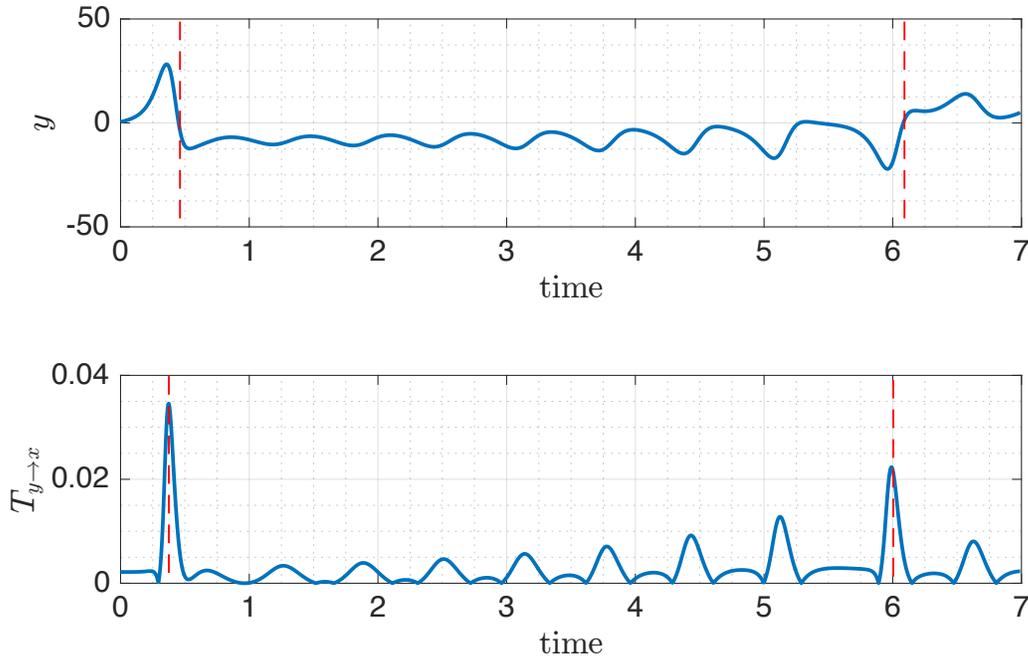
Figure 8.11   Information transfer and multiple phase transitions in Lorenz attractor.

8.5   Complex Networks

Complex networks are ubiquitous in nature, ranging from world wide web, social network (75; 84), to metabolic pathways of many organisms (85) etc. and phase transitions in such complex networks can be of different types like the emergence of giant connected components and condensation of edges. For example, consider a Erdos-Reyni network (86), with average degree $d(G)$. When $d(G) < 1$ there are a large number of isolated components and as $d(G) \to 1$, more and more isolated components start to merge together. Finally, when $d(G) > 1$, there is a single connected component, called the giant component. This phase transition, where a giant component is formed, is known as percolation (75).

In this preliminary study, we discuss the phase transition originating from the condensation of edges in simple graphs and demonstrate how the information transfer measure can identify the phase transitions.

Consider a linear dynamical system

$$
x(t+1) = \begin{pmatrix}
0 & 1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 1 & 0
\end{pmatrix} x(t) + \sigma\xi(t) \tag{8.8}
$$

where $x(t) \in \mathbb{R}^6$, $\sigma > 0$ is a small positive constant and $\xi(t)$ is independent and identical Gaussian noise of unit variance. The above dynamical system can be represented by a graph, as shown in Fig. 8.12.
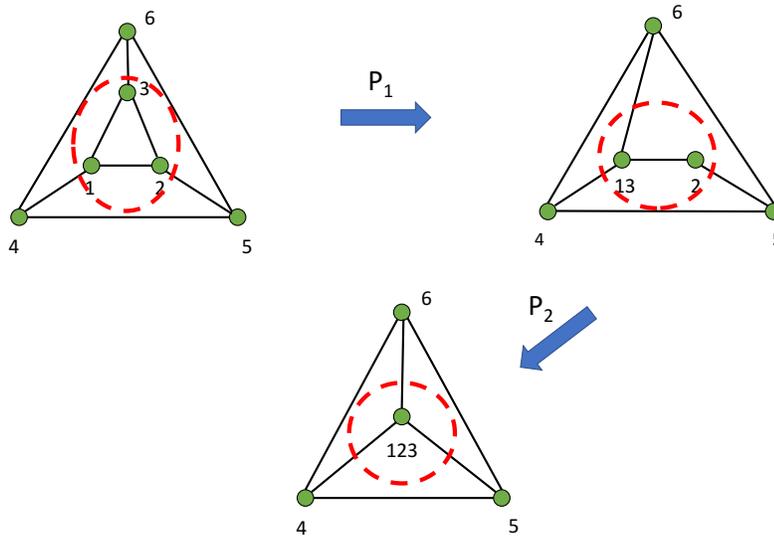
Figure 8.12   Phase transitions in a 6 node network.

The adjacency matrix of this graph is same as the system matrix. Note that, for a directed graph, the system matrix is the transpose of the adjacency matrix. We consider two phase transitions. The first transition (P_1) occurs, when the nodes 1 and 3 merge together to form the node 13 and the second transition (P_2) occurs when the node 13 and 2 merge together to form a single node (node 123). The phase transitions are shown in Fig. 8.12.
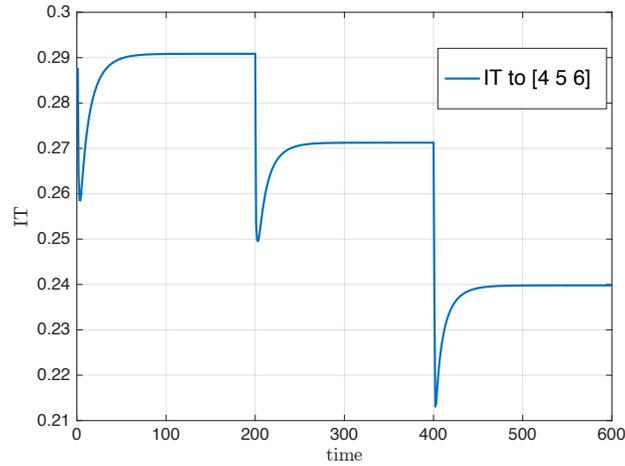
Figure 8.13   Information transfer in the 6 node network.

For simulating the phase transition, we simulated the original system for 200 seconds and after that nodes, 1 and 3 merged so that the first phase transition (P_1) occurs. We consider the information transfer to the nodes which do not collapse (in this case nodes $4, 5$ and 6) from the rest of the nodes. We find that at 200th second, the information transfer shows a sudden jump and not only that, it settles to a different value as well. Similar observations can be made when P_2 occurs. The jumps in the information transfer identify the phase transitions in the network and thus information transfer acts as an indicator for phase transition.

We discuss a second example, an exponential graph (86; 87; 88) of 20 nodes and consider two phase transitions as shown in Fig. 8.14.

In this case, the nodes 6 and 16 (marked by red in Fig. 8.14) collapse and become one single node during the first phase transition and during the second transition nodes 5 and 13 (marked by green in Fig. 8.14) collapse and become one single node. The first transition happens after the original system evolves for 200 seconds and the second transition happens after 400 seconds.

The networks that result from the phase transitions is shown in Fig. 8.15(a) and information transfer from the collapsed nodes to the rest of the network is plotted in Fig. 8.15(b). The information transfer again shows sudden jumps, identifying the phase transitions.

In these two examples, we considered the case of phase transition which results from the collapse of edges. We believe that exactly similar trends will be observed for the phase transition due to
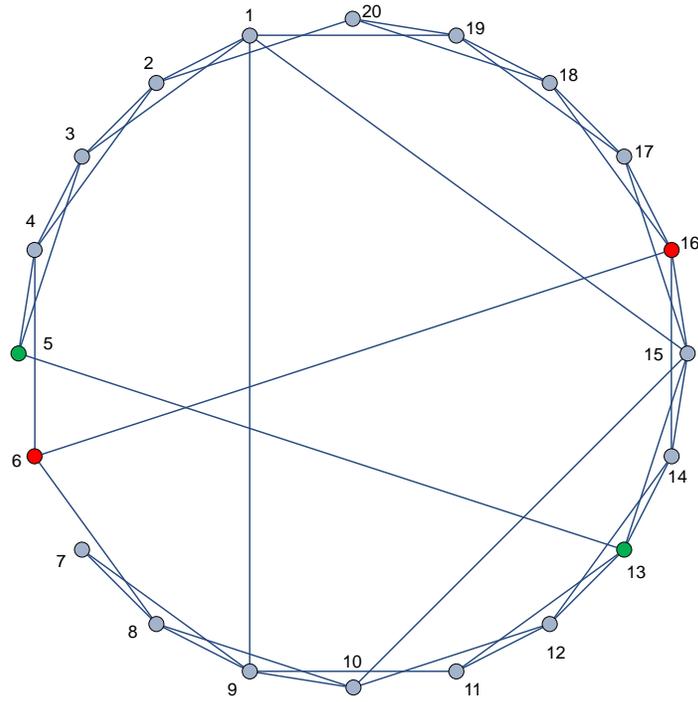
Figure 8.14   Exponential network of 20 nodes.

percolation. Analysis of percolation in terms of information transfer is left for future investigations as is the problem of working with real-life examples.

(a)

(b)
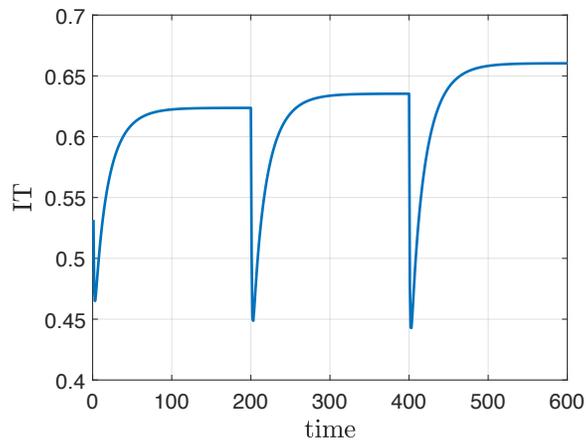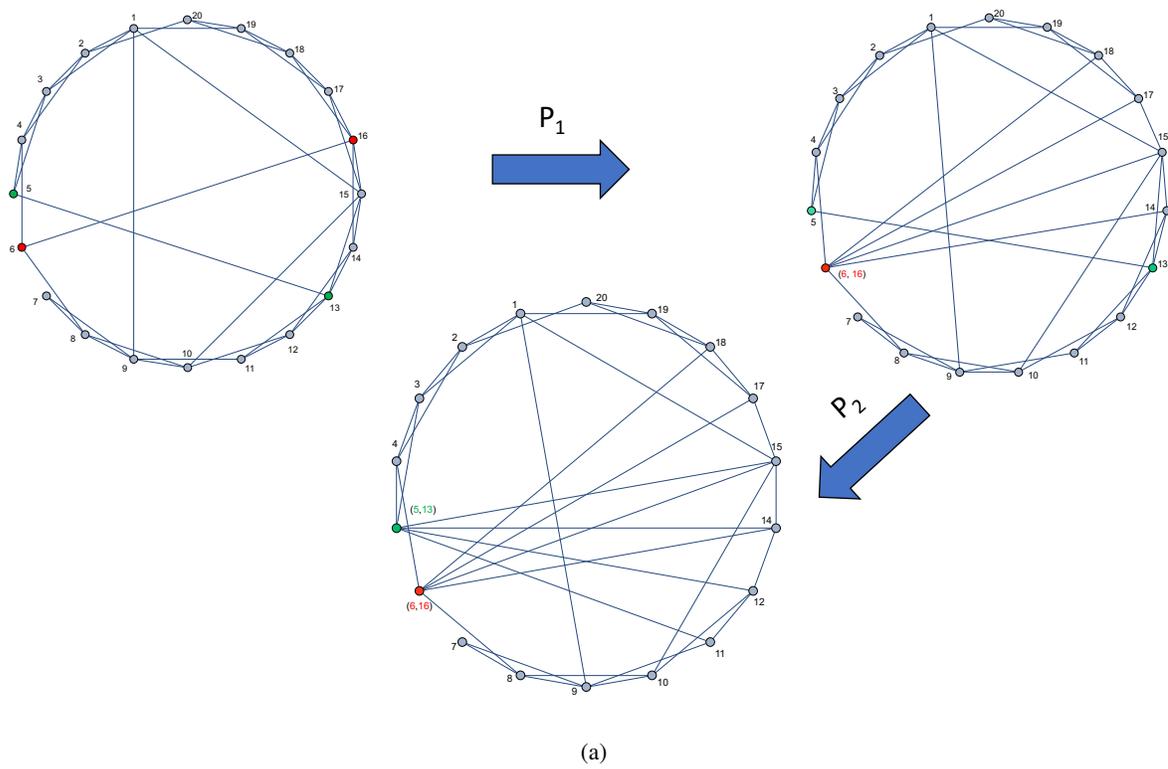
Figure 8.15 (a) Phase transitions in the exponential network. (b) Information transfer in the 20
node network.

CHAPTER 9.   CAUSAL INFERENCE FROM TIME SERIES DATA

So far, the formulation of information transfer was based on the assumption that one knows the underlying model of the dynamical system. However, in real life, in many different scenarios, one does not have access to the underlying mathematical model, but instead have the time series data from a dynamical system. In this chapter, we discuss the method of computation of one step information transfer from time series data. We assume that we can measure all the states of the system and use provide two algorithms for computing the one step information transfer. In particular, first we describe an algorithm which can be used on data coming from linear dynamical system, and later we describe an algorithm for computing information transfer for any general dynamical system, linear or non-linear.

9.1   Transfer operators for stochastic system

Consider a discrete time random dynamical system of the form

$$x_{t+1} = T(x_t, \xi_t) \tag{9.1}$$

where $T : X \times W \to X$ with $X \subset \mathbb{R}^N$ is assumed to be invertible with respect to x for each fixed value of ξ and smooth diffeomorphism. $\xi_t \in W$ is assumed to be independent identically distributed (i.i.d) random variable drawn from probability distribution ϑ i.e.,

$$\text{Prob}(\xi_t \in B) = \vartheta(B)$$

for every set $B \subset W$ and all t. Furthermore, we denote by $\mathcal{B}(X)$ the Borel-σ algebra on X and $\mathcal{M}(X)$ the vector space of bounded complex-valued measure on X. Associated with this discrete time dynamical system are two linear operators namely Koopman and Perron-Frobenius (P-F) operator. These two operators are defined as follows.

Definition 29 (Perron-Frobenius Operator). $\mathbb{P} : \mathcal{M}(X) \to \mathcal{M}(X)$ *is given by*

$$[\mathbb{P}\mu](A) = \int_{\mathcal{X}} \int_{W} \chi_A(T(x,v)) d\vartheta(v) d\mu(x) = \int_X p(x,A) d\mu(x) \tag{9.2}$$

where $\chi_A(x)$ is the indicator function for set A and $p(x,A)$ is the transition probability function.

For deterministic dynamical system $p(x,A) = \delta_{T(x)}(A)$. Under the assumption that $p(x,\cdot)$ is absolutely continuous with respect to Lebesgue measure, m, we can write

$$p(x,A) = \int_A k(x,y) dm(y)$$

for all $A \subset X$. Under this absolutely continuous assumption, the P-F operator on the space of densities $L_1(X)$ can be written as [1]

$$[\mathbb{P}g](y) = \int_X k(x,y) g(x) dm(x)$$

Definition 30 (Invariant measures). *Invariant measures are the fixed points of the P-F operator $\mathbb{P}$ that are additionally probability measures. Let $\bar\mu$ be the invariant measure then, $\bar\mu$ satisfies*

$$\mathbb{P}\bar\mu = \bar\mu$$

Under the assumption that the state space X is compact, it is known that the P-F operator admits at least one invariant measure.

Definition 31 (Koopman Operator). *Given any $h \in \mathcal{F}$, $\mathbb{U} : \mathcal{F} \to \mathcal{F}$ is defined by*

$$[\mathbb{U}h](x) = \mathbf{E}_\xi[h(T(x,\xi))] = \int_W h(T(x,v)) d\vartheta(v)$$

Properties 32. *Following properties for the Koopman and Perron-Frobenius operators can be stated.*

a). *For any function $h \in \mathcal{F}$ such that $h \geq 0$, we have $[\mathbb{U}h](x) \geq 0$ and hence Koopman is a positive operator.*

---

[1]with some abuse of notation we are using the same notation for the P-F operator defined on the space of measure and densities.

d). *If we define P-F operator act on the space of densities i.e., $L_1(X)$ and Koopman operator on space of $L_\infty(X)$ functions, then it can be shown that the P-F and Koopman operators are dual to each others as follows*

$$\langle \mathbb{U}f, g \rangle = \langle f, \mathbb{P}g \rangle$$

where $f \in L_\infty(X)$ and $g \in L_1(X)$.

e). *For $g(x) \geq 0$, $[\mathbb{P}g](x) \geq 0$.*

f). *Let $(X, \mathcal{B}, \mu)$ be the measure space where μ is a positive but not necessarily the invariant measure, then the P-F operator satisfies following property.*

$$\int_X [\mathbb{P}g](x) d\mu(x) = \int_X g(x) d\mu(x)$$

9.2 Robust approximation of P-F and Koopman operators

In this section, we derive the robust version of Extended Dynamic Mode Decomposition. Results for the robust implementation of DMD, Kernel EDMD and NSDMD will follow exactly along similar lines. Consider snapshots of data set obtained from simulating a discrete time random dynamical system $x \to T(x, \xi)$ or from an experiment

$$X = [x_0, x_2, \ldots, x_M] \tag{9.3}$$

where $x_i \in X \subset \mathbb{R}^n$. The data-set $\{x_k\}$ can be viewed as sample path trajectory generated by random dynamical system and could be corrupted by either process or measurement noise or both. A large number of sample path trajectories need to be simulated to realize sufficient statistics of the random dynamical system. However, in practice, only few sample path trajectories over finite time horizon are available, and it is hard to approximate the statistics of RDS using the limited amount of data-set. Furthermore, rarely one knows the probability distribution of the underlying noise process, i.e., ϑ. Estimating ϑ is in itself a challenging problem. In spite of these difficulties, it is essential to develop an algorithm for the approximation of transfer operators that explicitly account for the uncertainty in

data-set. We propose a robust optimization-based approach to address this challenge. In particular, we consider deterministic, but norm bounded uncertainty in the data set. Since the trajectory $\{x_k\}$ is one particular realization of the RDS, the other random realization can be assumed to be obtained by perturbing $\{x_k\}$. We assume that the data points x_k are perturbed by norm bounded deterministic perturbation of the form

$$\delta x_k = x_k + \delta, \quad \delta \in \Delta.$$

Several possible choices for the uncertainty set Δ can be considered. For example

$$\Delta := \{\delta \in \mathbb{R}^n : \parallel \delta \parallel_2 \le \rho\} \tag{9.4}$$

restrict the 2-norm of δ to ρ. Another possible choice could be

$$\Delta := \{\delta \in \mathbb{R}^n : \parallel \delta \parallel_{Q_i} \le 1, \ i = 1, \ldots, d\} \tag{9.5}$$

where $Q_i \ge 0$ and implies that uncertainty δ lies at the intersection of ellipsoids. More generally, one can also consider Δ set to be of the form

$$\Delta = \{\delta \in \mathbb{R}^n : h_i(\delta) \le 0, \ i = 1, \ldots, d\} \tag{9.6}$$

for some convex function $h_i(\delta)$. These different choices for the uncertainty set Δ allow us to encapsulate the information about the uncertainty δ. For example, if the vector random variable in $\mathbb{R}^2$ is uniformly distributed with support in interval $[-d, d] \times [-d, d]$, then the uncertainty set Δ can be written as intersection of two ellipsoids $\mathcal{E}_1 = \{(x, y) | \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1\}$ and $\mathcal{E}_2 = \{(x, y) | \frac{x^2}{b^2} + \frac{y^2}{a^2} \le 1\}$ where $a >> d$ and $1/a^2 + 1/b^2 = 1/d^2$(Fig. 9.1(a)). For example, let $d = 0.5$. Then $[-0.5, 0.5] \times [-0.5, 0.5]$ can be expressed as the intersection of the two ellipsoids $\mathcal{E}_1$ and $\mathcal{E}_2$, with $a = 50$ and $b = 0.4996$, as shown in figure 9.1(b).

Now let $\mathcal{D} = \{\psi_1, \psi_2, \ldots, \psi_K\}$ be the set of dictionary functions or observables. The dictionary functions are assumed to belong to $\psi_i \in L_2(X, \mathcal{B}, \mu) = \mathcal{G}$, where μ is some positive measure, not necessarily the invariant measure of T. Let $\mathcal{G}_\mathcal{D}$ denote the span of $\mathcal{D}$ such that $\mathcal{G}_\mathcal{D} \subset \mathcal{G}$. The choice of dictionary functions are very crucial and it should be rich enough to approximate the leading eigenfunctions of Koopman operator. Define vector valued function $\boldsymbol{\Psi} : X \to \mathbb{C}^K$ as

$$\boldsymbol{\Psi}(x) := \begin{bmatrix} \psi_1(x) & \psi_2(x) & \cdots & \psi_K(x) \end{bmatrix} \tag{9.7}$$
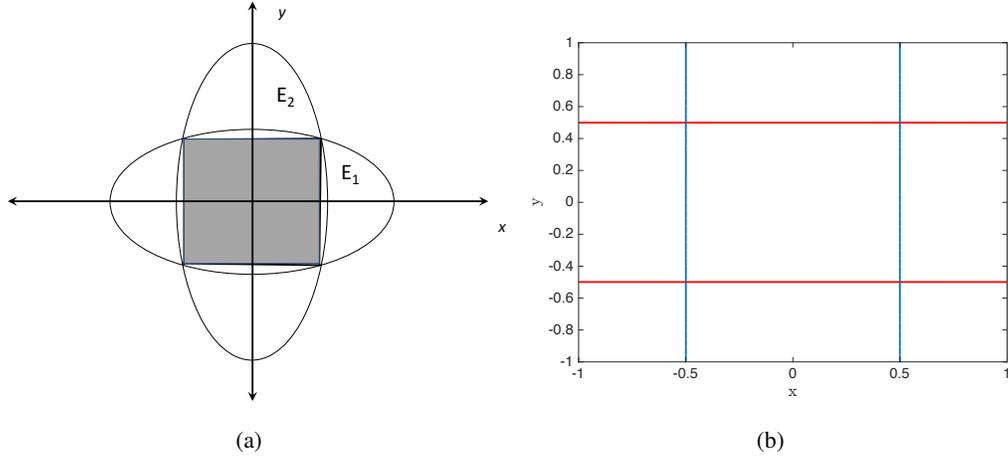
Figure 9.1 (a) Uncertainty set as intersection of two ellipsoids. (b) Representation of the uncertainty set $\Delta = [-0.5, 0.5] \times [-0.5, 0.5]$.

In this application, $\boldsymbol{\Psi}$ is the mapping from physical space to feature space. Any function $\phi, \hat{\phi} \in \mathcal{G}_\mathcal{D}$ can be written as

$$\phi = \sum_{k=1}^{K} a_k \psi_k = \boldsymbol{\Psi}\boldsymbol{a}, \quad \hat{\phi} = \sum_{k=1}^{K} \hat{a}_k \psi_k = \boldsymbol{\Psi}\hat{\boldsymbol{a}} \tag{9.8}$$

for some set of coefficients $\boldsymbol{a}, \hat{\boldsymbol{a}} \in \mathbb{C}^K$. Let

$$\hat{\phi}(x) = [\mathbb{U}\phi](x) + r = E_\xi[\phi(T(x,\xi))] + r. \tag{9.9}$$

Unlike deterministic case where we evaluate (9.9) at the data point $\{x_k\}$, for the uncertain case we do not have sufficient data points to evaluate the expected value in the above expression. Instead we use the fact that different realizations of the RDS will consist of the form $\{x_k + \delta\}$ with $\delta \in \Delta$ to write (9.9) as follows:

$$\hat{\phi}(x_m + \delta x_m) = \phi(x_{m+1}) + r, \quad k = 1, \ldots, M - 1. \tag{9.10}$$

The objective is to minimize the residual for not just one pair of data points $\{x_m, x_{m+1}\}$, but over all possible pairs of data points of the form $\{x_m + \delta, x_{m+1}\}$. Using (9.8) we write the above as follows:

$$\boldsymbol{\Psi}(x_k + \delta x_k)\hat{\boldsymbol{a}} = \boldsymbol{\Psi}(x_{k+1})\boldsymbol{a} + r.$$

We seek to find matrix $\mathbf{K}$, the finite dimensional approximation of Koopman operator that maps coefficient vector $\boldsymbol{a}$ to $\hat{\boldsymbol{a}}$, i.e., $\mathbf{K}\boldsymbol{a} = \hat{\boldsymbol{a}}$, while minimizing the residual term, r. Premultiplying by

$\mathbf{\Psi}^{\top}(x_m)$ on both the sides of above expression and summing over m we obtain

$$\left[\frac{1}{M} \sum_m \mathbf{\Psi}^{\top}(x_m)\mathbf{\Psi}(x_m + \delta x_m)\mathbf{K} - \mathbf{\Psi}^{\top}(x_m)\mathbf{\Psi}(x_{m+1}) \right] \boldsymbol{a}.$$

In the absence of the uncertainty the objective is to minimize the appropriate norm of the quantity inside the bracket over all possible choices of matrix $\mathbf{K}$. However, for robust approximation, presence of uncertainty acts as an adversary whose goal is to maximize the residual term. Hence the robust optimization problem can be formulated as a $\min - \max$ optimization problem as follows.

$$\min_{\mathbf{K}} \max_{\delta \in \Delta} \| \mathbf{G}_\delta \mathbf{K} - \mathbf{A} \|_F =: \min_{\mathbf{K}} \max_{\delta \in \Delta} \mathcal{F}(\mathbf{K}, \mathbf{G}_\delta, \mathbf{A}) \tag{9.11}$$

where

$$\mathbf{G}_\delta = \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Psi}(x_m)^{\top}\mathbf{\Psi}(x_m + \delta x_m)$$

$$\mathbf{A} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Psi}(x_m)^{\top}\mathbf{\Psi}(x_{m+1}), \tag{9.12}$$

with $\mathbf{K}, \mathbf{G}_\delta, \mathbf{A} \in \mathbb{C}^{K \times K}$. The $\min - \max$ optimization problem (9.11) is in general nonconvex and will depend on the choice of dictionary functions. This is true because $\mathcal{F}$ in (9.11) is not in general concave function of δ for fixed $\mathbf{K}$. Hence, we convexify the problem as follows

$$\min_{\mathbf{K}} \max_{\delta \mathbf{G} \in \bar{\Delta}} \| (\mathbf{G} + \delta \mathbf{G})\mathbf{K} - \mathbf{A} \|_F \tag{9.13}$$

where $\delta \mathbf{G} \in \mathbb{R}^{K \times K}$ is the new perturbation term characterized by uncertainty set $\bar{\Delta}$ which lies in the feature space of dictionary function and the matrix $\mathbf{G} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Psi}(x_m)^{\top}\mathbf{\Psi}(x_m)$. $\bar{\Delta}$ is the new uncertainty set defined in the feature space and will inherit the structure from set Δ in the data space. In particular, following result can be used to connect the two uncertainty set, Δ and $\bar{\Delta}$ when Δ is described as in Eq. (9.4).

Proposition 33. *In the convex problem (9.13), δG is bounded as*

$$\| \delta G \|_F \leq \lambda \Lambda \Gamma \tag{9.14}$$

where $\| \delta x_m \|_F \leq \lambda$, $\| \mathbf{\Psi}(x_m) \|_F \leq \Lambda$ and $\| \mathbf{\Psi}'(x_m) \|_F \leq \Gamma$ for all m.

Proof. From Taylor series expansion we have, $\mathbf{\Psi}(x_m + \delta x_m) = \mathbf{\Psi}(x_m) + \mathbf{\Psi}'(x_m)\delta x_m$, where $\mathbf{\Psi}'(x_m)$ is the first derivative of $\mathbf{\Psi}(x)$ at x_m. Hence,

$$G_\delta = G + \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Psi}^\top(x_m)\delta x_m \mathbf{\Psi}'(x_m)$$

Let $\delta G = \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Psi}^\top(x_m)\delta x_m \mathbf{\Psi}'(x_m)$. Hence,

$$
\begin{aligned}
\| \delta G \|_F &= \| \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Psi}^\top(x_m)\delta x_m \mathbf{\Psi}'(x_m) \|_F \\
&\leq \frac{1}{M} \sum_{m=1}^{M} \| \mathbf{\Psi}^\top(x_m)\delta x_m \mathbf{\Psi}'(x_m) \|_F \\
&\leq \frac{1}{M} \sum_{m=1}^{M} \| \mathbf{\Psi}^\top(x_m) \|_F \cdot \| \delta x_m \|_F \cdot \| \mathbf{\Psi}'(x_m) \|_F \\
&\leq \lambda \Lambda \Gamma
\end{aligned}
$$

$\square$

Similar results can be used to connect the two uncertainty set Δ and $\bar{\Delta}$ for the case where Δ is described in Eqs. (9.5) and (9.6) to convert non-convex optimization problem to convex. For the uncertainty set given in Eq. (9.5), it can be easily shown that the problem simplifies to case described in the above proposition, where λ depends on the eigenvalues of the matrices Q_i. Similarly, for the uncertainty set given in (9.6), under the assumption that Δ is compact, it also boils down to proposition 33.

In (89), we proposed Naturally Structured Dynamic Mode Decomposition (NSDMD) algorithm for finite dimensional approximation of the transfer Koopman and P-F operator. Apart from preserving positivity and Markov properties of the transfer operator, this algorithm exploits the duality between P-F and Koopman operator to provide the approximation of P-F operator. The algorithm presented for the robust approximation of Koopman operator can be combined with NSDMD for the robust approximation of P-F operator. In particular, under the assumption that all the dictionary functions are positive, following modification can be made to optimization formulation (9.13) for the approximation of Koopman operator.

$$\min_{\mathbf{K}} \max_{\delta\mathbf{G}\in\bar{\Delta}} \| (\mathbf{G} + \delta\mathbf{G})\mathbf{K} - \mathbf{A} \|_F$$

$$\text{s.t.} \quad \mathbf{K}_{ij} \geq 0$$

$$[\Lambda\mathbf{K}\Lambda^{-1}]_{ij} \geq 0$$

$$\Lambda\mathbf{K}\Lambda^{-1}\mathbb{1} = \mathbb{1} \tag{9.15}$$

where $\Lambda = \langle \boldsymbol{\Psi}(x), \boldsymbol{\Psi}(x) \rangle$ with $[\Lambda]_{ij} = \langle \psi_i, \psi_j \rangle$ is symmetric positive definite matrix. We refer the interested reader to (89) for details of NSDMD formulation. Using duality the robust approximation of the P-F operator, $\mathbf{P}$, can then be written as $\mathbf{P} = \mathbf{K}^\top$. Most common approach for solving the robust optimization problem is by using a robust counterpart. In the following section we show that the robust counterpart of the robust optimization problem can be constructed and is a convex optimization problem.

9.2.1 Robust Optimization, Regularization, and Sparsity

The robust optimization problem (9.13) has some interesting connection with optimization problems involving regularization term. In particular, we have following Theorem.

Theorem 34. *Following two optimization problems*

$$\min_{\mathbf{K}} \max_{\delta\mathbf{G}:\|\delta\mathbf{G}\|_F \leq \lambda} \| (\mathbf{G} + \delta\mathbf{G})\mathbf{K} - \mathbf{A} \|_F \tag{9.16}$$

$$\min_{\mathbf{K}} \| \mathbf{G}\mathbf{K} - \mathbf{A} \|_F + \lambda \| \mathbf{K} \|_F \tag{9.17}$$

are equivalent.

Proof. The following proof is from (90), and we outline here for the convenience of the reader. For any $\delta\mathbf{G} \in \bar{\Delta}$, such that $\| \delta\mathbf{G} \|_F \leq \lambda$, $\| (\mathbf{G} + \delta\mathbf{G})\mathbf{K} - \mathbf{A} \|_F = \| (\mathbf{G}\mathbf{K} + \delta\mathbf{G}\mathbf{K} - \mathbf{A}) \|_F$. By triangle inequality, using the fact that $\| \delta\mathbf{G} \|_F \leq \lambda$, we have $\| (\mathbf{G} + \delta\mathbf{G})\mathbf{K} - \mathbf{A} \|_F \leq \| \mathbf{G}\mathbf{K} - \mathbf{A} \|_F + \lambda \| \mathbf{K} \|_F$. On the other hand, given any $\mathbf{K}$, we can choose a $\delta\mathbf{G}$, such that $\delta\mathbf{G}\mathbf{K}$ is aligned with $(\mathbf{G}\mathbf{K} - \mathbf{A})$ and thus the two problems are equivalent. $\qquad\square$

The regularization term penalize the Frobenius norm of the matrix $\mathbf{K}$ and is also known as Tikhonov regularization. Equivalence to the more popular ℓ_1 regularization or Lasso regularization can also be shown if we change the structure of the uncertainty set. Note that in the above theorem the uncertainty set $\bar{\Delta}$ is defined as follows:

$$\bar{\Delta} := \{\delta\mathbf{G} \in \mathbb{R}^{K \times K} : \| \delta\mathbf{G} \| \leq \lambda\}$$

The equivalence between the robust optimization problem (9.13) and the ℓ_1 Lasso regularization can be established as follows.

Theorem 35. *Define*

$$\bar{\Delta} := \{\delta\mathbf{G} = (\delta\mathbf{G}_1, \ldots, \delta\mathbf{G}_K) \in \mathbb{R}^{K \times K} : \| \delta\mathbf{G}_i \|_2 \leq c\}.$$

Following two optimization problems are equivalent

$$\min_{\mathbf{K}} \max_{\delta\mathbf{G} \in \bar{\Delta}} \| (\mathbf{G} + \delta\mathbf{G})\mathbf{K} - \mathbf{A} \|_F \tag{9.18}$$

$$\min_{\mathbf{K}} \| \mathbf{GK} - \mathbf{A} \|_F + c\sum_{k=1}^{K} \| \mathbf{K}_k \|_1 \tag{9.19}$$

where $\mathbf{K}_k$ is the kth column of matrix $\mathbf{K}$.

Refer to (90) for the proof. It is well known that the optimization problem with the regularization term especially the ℓ_1 Lasso type regularization induce sparsity structure on the optimal solution. The equivalence between robust optimization and regularized optimization problem provides an alternate point of view to the sparsity structure, i.e., *robustness implies sparsity*. In (91), sparse DMD algorithm is proposed where ℓ_1 type regularization term is used to impose sparsity structure on the amplitude terms which appear in the temporal expansion of data along dynamic modes. We expect similar robust optimization-based viewpoint can be provided to the sparse DMD algorithm proposed in (91).

9.2.2 Design of robust predictor

While the optimization problem with the regularization term (9.17) and the robust optimization problem (9.16) are equivalent, the robust optimization point of view to the optimization problem

offers a particular advantage. In particular, robust optimization viewpoint provides a systematic way of determining the regularization parameter which often is a tuning parameter. On the other hand, the optimization formulation with the regularization term has interesting interpretation borrowed from machine learning literature. In problems involving model fitting from data one of the fundamental tradeoff arise between the quality of approximation and complexity of approximation function. For example, in the absence of the regularization term, the optimization problem will find a matrix $\mathbf{K}$ that best tries to fit the training data to the model. Hence, without the regularization term, there will be a tendency to over-fit model parameters to data. Such over-fitted model will perform well on the training data and give a smaller value of optimal cost. However, these over-fitted models will often perform poorly on the test data-set. On the other hand, the optimization problem with the regularization term tries to strike a balance between over fitting and prediction. This observation on the role of regularization term has an important implication on the proposed application of transfer operator framework for the design of data-driven predictor. For the design of predictor dynamics, we first use training data for the approximation of the transfer Koopman operator. Let $\{x_0, \ldots, x_M\}$ be the training data-set and $\mathbf{K}$ be the finite-dimensional approximation of the transfer Koopman operator obtained using the robust algorithm (9.13). Let $\bar{x}_0$ be the initial condition from which we want to predict the future. The initial condition from state space is mapped to the feature space using the same choice of basis function used in the robust approximation of Koopman operator i.e.,

$$\bar{x}_0 \implies \boldsymbol{\Psi}(\bar{x}_0)^\top =: \mathbf{z} \in \mathbb{R}^K.$$

This initial condition is propagated using Koopman operator as

$$\mathbf{z}_n = \mathbf{K}^n \mathbf{z}.$$

The predicted trajectory in the state space is then obtained as

$$\bar{x}_n = C \mathbf{z}_n$$

where matrix C is obtained as the solution of following least square problem

$$\min_C \sum_{i=1}^M \| x_i - C \boldsymbol{\Psi}(x_i) \|_2^2 \tag{9.20}$$

In the simulation section, we demonstrate the effectiveness of the proposed robust prediction algorithm on linear and nonlinear systems.

9.2.3 Examples

Example 36. *System of Coupled Oscillators :*

We consider a network of linear coupled oscillators forced with white Gaussian noise process

$$\ddot{\theta}_k = -\mathcal{L}_k\theta - d\dot{\theta}_k + \sigma\frac{d\xi}{dt}, \quad k = 1, \cdots, N \tag{9.21}$$

where θ_k is the angular position of the kth oscillator, N is the number of oscillators, $\mathcal{L}_k$ is the kth row of the Laplacian $\mathcal{L}$. The Laplacian corresponds to the star network and is given as follows:

$$\mathcal{L} = \begin{pmatrix} 7 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

d is the damping coefficient.

For simulation purpose, we took $N = 8$ and $d = 0.05$ and $\sigma = 0.4$. We assume that the output is the angular position of all the oscillators and is not subjected to noise disturbance. The system was discretized with $\delta t = 0.01$ and data was collected for 60 times steps. However, among the 60 data points, only the first 20 data points were used as the training data for approximating the Koopman operator, using both the proposed robust algorithm and the non-robust algorithm. Since the system is linear, we use linear dictionary function for the approximation of Koopman operator. In Fig. (9.2)a, we show the first few dominant eigenvalues obtained using the original DMD algorithm, i.e., DMD algorithm applied to data-set without noise (clean data-set), proposed robust DMD algorithm, and regular DMD applied to noisy data-set. The regularization parameter for robust DMD algorithm was chosen to be 0.4. From Fig. (9.2)a, we notice that the robust DMD does a better job in approximating the dominant eigenvalues of the original noise-free system compared to regular DMD. Furthermore,

the regular DMD also leads to unstable eigenvalue which is not present in the original system since the original system has positive damping.
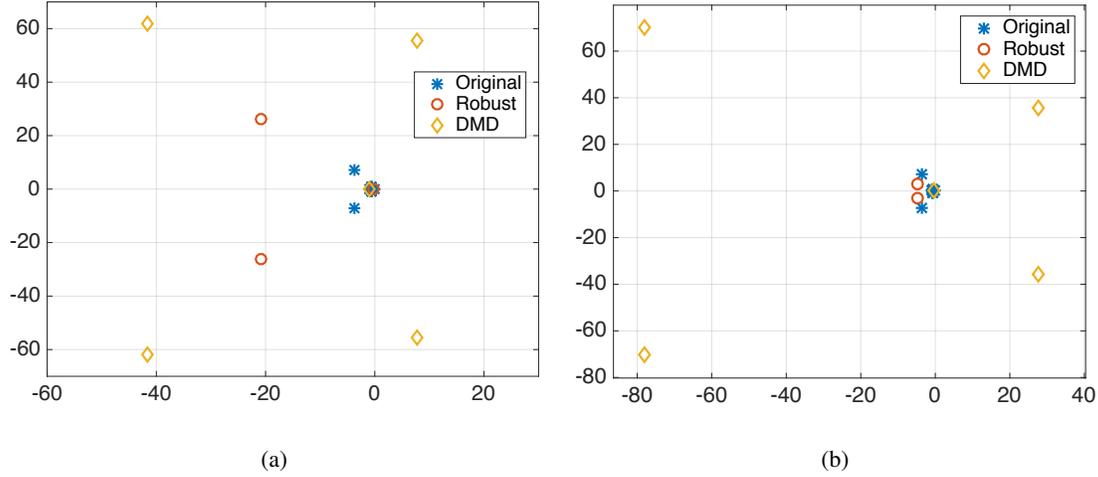
Figure 9.2   (a) Dominant eigenvalue comparison when there is no measurement noise. (b) Dominant eigenvalue comparison with both process and measurement noise.

The second set of simulations in Fig. (9.2)b are obtained with both measurement and process noise. Both the measurement and process noise are assumed white with variance of 0.3. The measurement data was again collected for 60 time steps, of which the first 20 were used as the training data. Eigenvalues computed using the regular DMD are further in right half plane whereas the Koopman eigenvalues obtained using robust DMD are all stable and also captures the first two dominant eigenvalues of the original system.

Example 37. *Noisy rotations on a circle*

We consider the example discussed in (92),(93). A dynamical system T, corresponding to rotation on the unit circle S^1 is given by

$$T(x) = x + \theta \tag{9.22}$$

where $\theta \in S^1$ is a constant number. We consider a stochastic dynamical system with process noise, so that the RDS is given by

$$T(x, \xi) = x + \theta + \xi \tag{9.23}$$

where ξ are independent and identically distributed random variables taking values in $[-0.7, 0.7]$. In this example we considered $x(0) = 1$, $\theta = \pi/320$. We consider the dictionary functions as

$$\mathbf{\Psi} = e^{2\pi inx} \tag{9.24}$$

where $n = -50, -49, \cdots, 49, 50$. Data was collected for 6000 time steps and data for the first 50 time steps was used for training purpose.

(a)

(b)

(c)
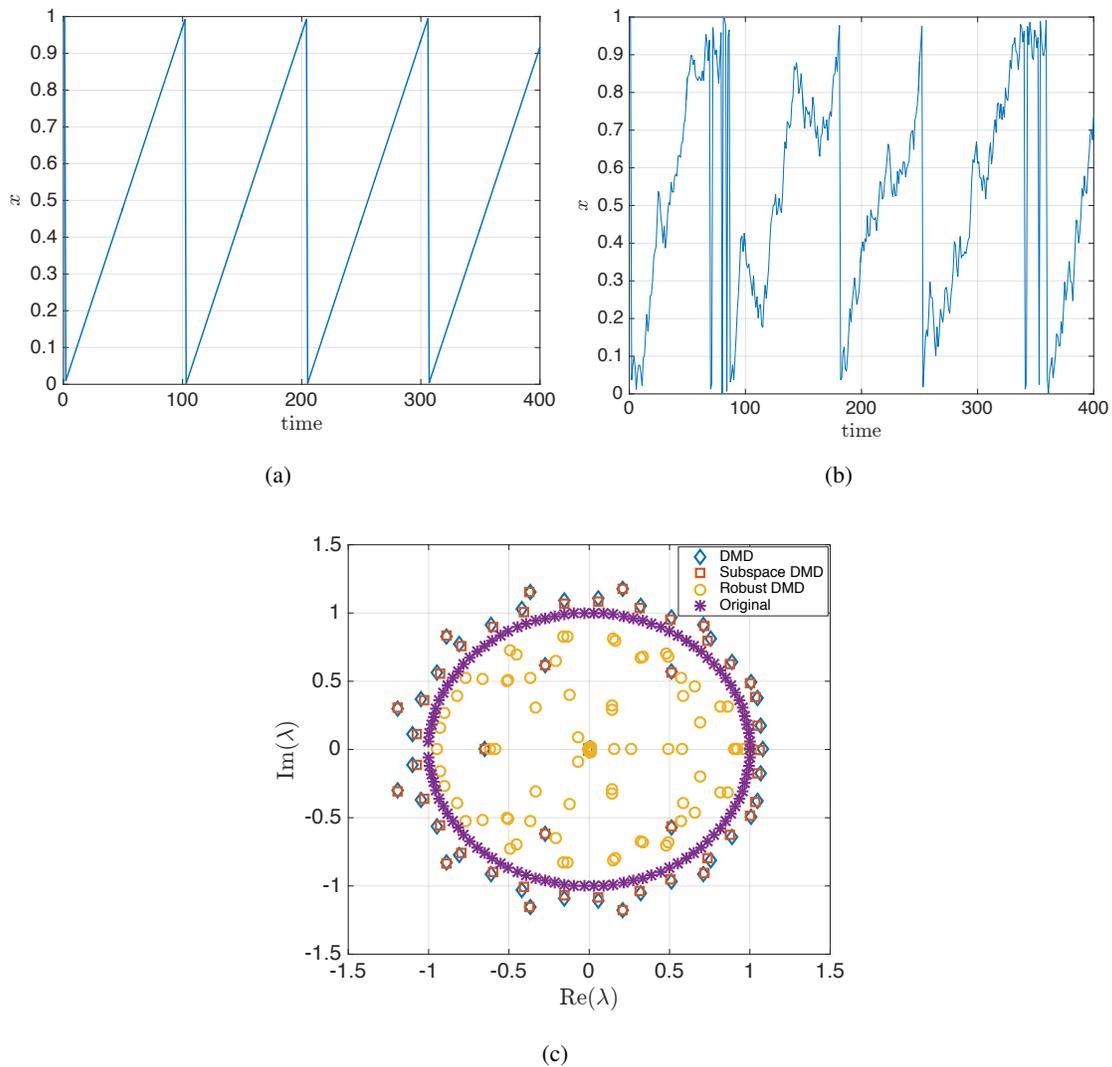
Figure 9.3 (a) Clean data. (b) Noisy data. (c) Comparison of eigenvalues obtained with different algorithms.

The clean data and the noisy data is shown in Fig. 9.3(a) and (b) respectively. The eigenvalues obtained using the different algorithms is plotted in Fig. 9.3(c) and it can be seen that the robust DMD

algorithm provides closer match for the eigenvalues of the deterministic system. Moreover, normal DMD and subspace DMD yields unstable eigenvalues, though the original system is not unstable. As mentioned earlier, normal DMD and subspace DMD works for RDS when the training data set is large enough, but in reality it may not be possible to obtain such large data sets and in that case robust DMD provides much better approximation of the Koopman operator.

Example 38. *The nonlinear stochastic Stuart-Landau system :*

The nonlinear stochastic Stuart-Landau system (94). The stochastic Stuart-Landau equation on a complex function $z(t) = r(t)\exp(i\theta(t))$ is given by

$$\dot{z} = (\mu + i\gamma)z - (1 + i\beta)|z|^2 z + \sigma\xi(t), \tag{9.25}$$

where $\xi(t)$ is a white Gaussian noise of unit variance and i is the imaginary unit. In absence of process noise, the solution of (9.25) evolves on the limit cycle $|z| = \sqrt{\mu}$. Hence, the continuous time eigenvalues lie on the imaginary axis, if the process noise is absent. The discretized version of (9.25) is

$$\begin{pmatrix} r_{t+1} \\ \theta_{t+1} \end{pmatrix} = \begin{pmatrix} r_t + (\mu r_t - r_t^3)\delta t \\ \theta_t + (\gamma - \beta r_t^2)\delta t \end{pmatrix} + \sigma_p \begin{pmatrix} \delta t & 0 \\ 0 & \delta t/r_t \end{pmatrix} \xi_t. \tag{9.26}$$

We assume that the observation are corrupted with noise and of the form

$$y_t = \begin{pmatrix} e^{-10i\theta_t} & e^{-9i\theta_t} & \cdots & e^{9i\theta_t} & e^{10i\theta_t} \end{pmatrix} + \sigma_o w_t. \tag{9.27}$$

The noisy output data was used to construct the finite dimensional approximation of the operator, with $\delta t = 0.01$. Both the process noise and measuement noise considered here are uniform with support $[-.03, 0.3]$ and $[-0.1, 0.1]$ respectively. A sample trajectory is shown in Fig. 9.4.

The system was initialized at $(1, -\pi)$ and data were obtained for 100 time steps and data of the first 30 steps was used as the training data and the prediction was made over future steps. The dictionary functions we chose are $e^{in\theta_t}$, $n = -10, -9, \cdots, 9, 10$. The eigenvalues of the approximate robust DMD Koopman operator, normal DMD Koopman and subspace DMD operator for the initial condition $(1, -\pi)$ are shown in Fig. 9.5(a). It is observed that all the algorithms are capturing the oscillatory nature of the underlying system, but normal DMD and subspace DMD generate a
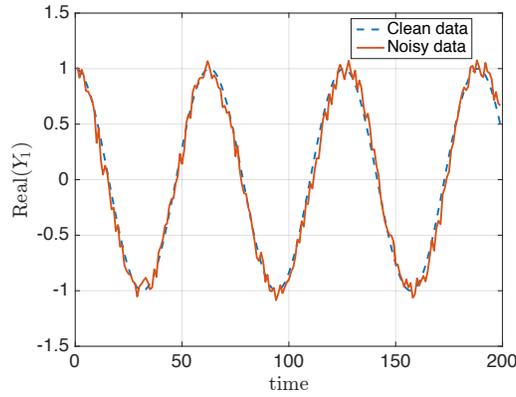
Figure 9.4   Sample output trajectory

(a)   (b)
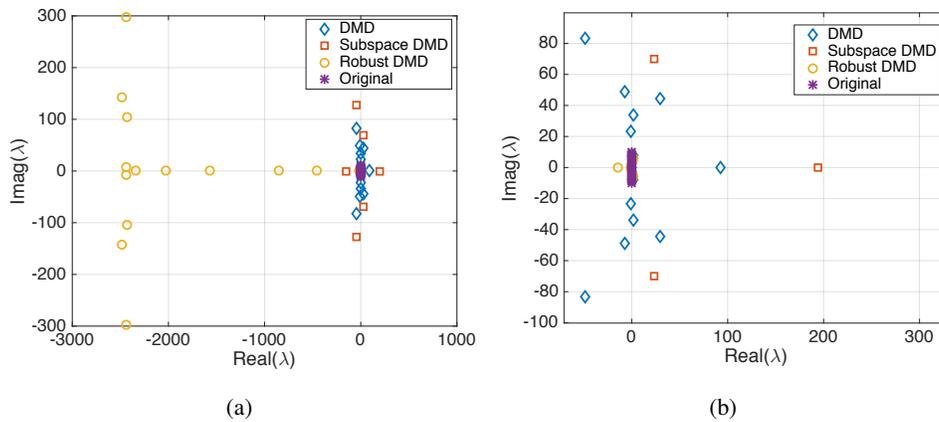
Figure 9.5   (a) Eigenvalue comparison of Robust DMD, normal DMD and subspace DMD. (b) Dominant eigenvalues.

(a)   (b)
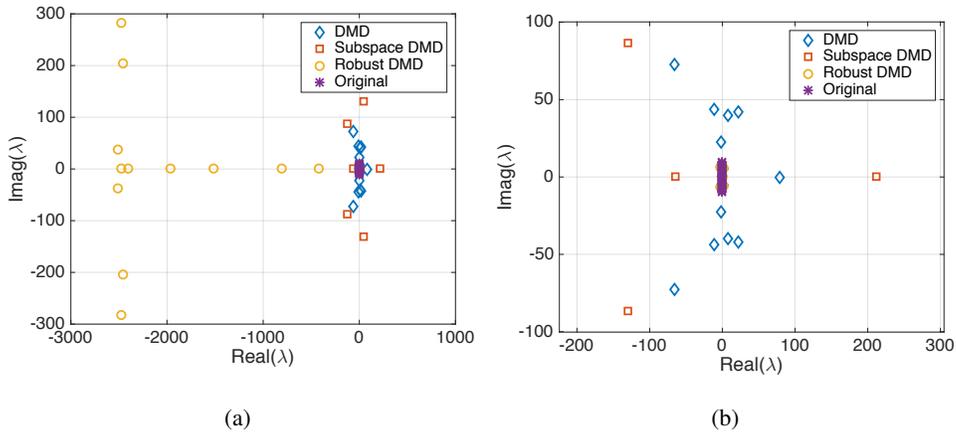
Figure 9.6   (a) Eigenvalue comparison of Robust DMD, normal DMD and subspace DMD. (b) Dominant eigenvalues.

Koopman operator which has an eigenvalue on the right half plane implying the system to be unstable. The first few dominant eigenvalues obtained using the three different algorithms are shown in Fig. 9.5(b).

(a)   (b)
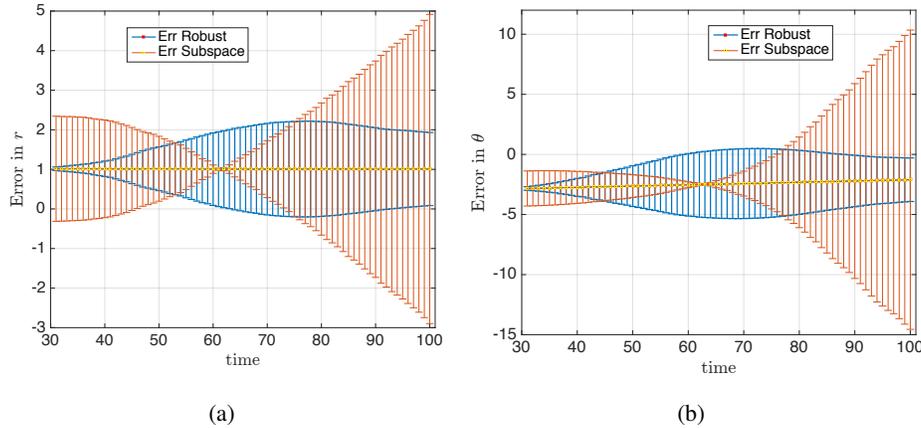
Figure 9.7   (a) Error in prediction in r for initial state at $(1, -\pi)$. (b) Error in prediction in θ initial state at $(1, -\pi)$.

Once the Koopman operators are obtained, one can use the obtained operator to predict the future. In Fig. 9.7(a) and (b), we compare the errors in prediction in r and θ respectively, when using robust DMD and subspace DMD for the system starting from $(1, -\pi)$. The errors are plotted around the actual values of r and θ. For example, the error in r is plotted around the actual value of $r = 1$. As mentioned earlier, the first 30 time step data was used as the training data, and the operators obtained using the training data has been used to predict future 70 time steps. As can be seen from Fig. 9.7(a) and (b), robust DMD algorithm yields much smaller error when compared to subspace DMD. The error in prediction using normal DMD increases exponentially and is not shown in the error plot. However, we find that the error using subspace DMD decreases initially and then grows exponentially. This is because the coupling from the unstable subspace to the stable subspace is small and hence initially the dominant stable eigenvalues make the error decrease but as the system evolves, the effect of the unstable eigenvalue becomes more prominent and the error grows exponentially.

As mentioned earlier, compared to subspace DMD Robust DMD provides a better approximation of systems, when the size of training data is not substantial. To illustrate this point, we used different sizes of training dataset and used it to predict future 10 time steps. In particular, we varied the training

data from 10 time steps to 40 time steps and looked at the average error in prediction in both r and θ

over 10 future time steps.
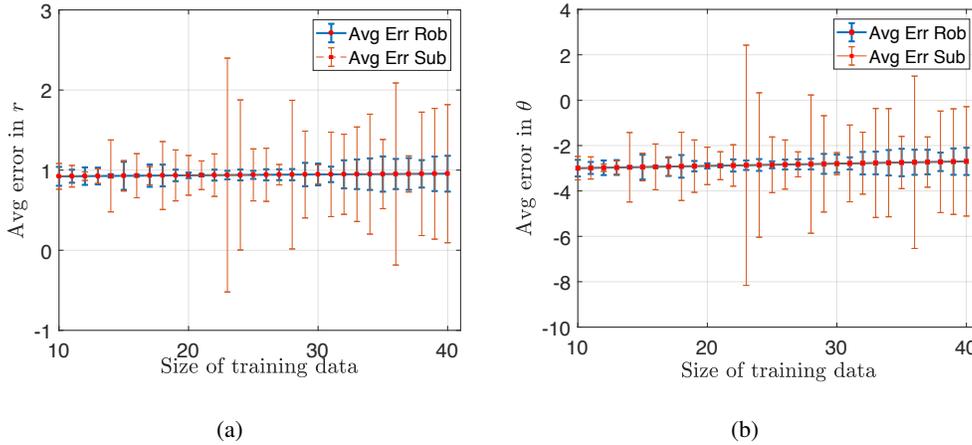
Figure 9.8   (a) Average error in prediction of r. (b) Average error in prediction of θ.

In Fig. 9.8(a) we plot the average error in prediction in r. The x-axis shows the number of time steps used for training and for each such training data set, the state r was predicted for 10 future time steps and the average error in the prediction is plotted along the y-axis. It can be observed that for all the different sizes of training data used, average error for Robust DMD algorithm always shows smaller error compared to subspace DMD. Similarly, in Fig. 9.8(b), we plot the average error is prediction of θ and observe the same.

Example 39. *Stochastic Burger equation*

The third example we consider in this paper is stochastic Burger's equation

$$\partial_t u(x,t) + u\partial_x u = k\partial_x^2 u + \sigma_p e(x,t)$$

with $k = 0.01$, $e(x,t)$ is uniform random distribution with support $[-1, 1]$ and $\sigma_p = 0.2$. In the simulation, we approximated the PDE solution using the Finite Difference method (95) with the initial condition $u(x,0) = sin(2\pi x)$ and Dirichet boundary condition $u(0,t) = u(1,t) = 0$. Given the spatial and temporal ranges, $x \in [0,1]$, $t \in [0,1]$, the discretizaion steps are chosen as $\Delta t = 0.02$ and $\Delta x = 1 \times 10^{-2}$.

The generated data is plotted in Fig. 9.9(a). The robust DMD algorithm, subspace DMD algorithm and the regular DMD algorithm are applied to a data set collected over 100 time steps. We assumed both process and measurement noise and each has a variance of 0.2. The eigenvalues obtained using regular DMD, subspace DMD and robust DMD approach are shown in Fig. 9.9)(b) and zoomed in plot shown in Fig. 9.9(c). As in the previous examples, we find that one of the eigenvalues obtained using regular DMD and subspace DMD approach is unstable. Furthermore, the dominant eigenvalues show a closer match for robust DMD compared to other methods.

(a)

(b)

(c)
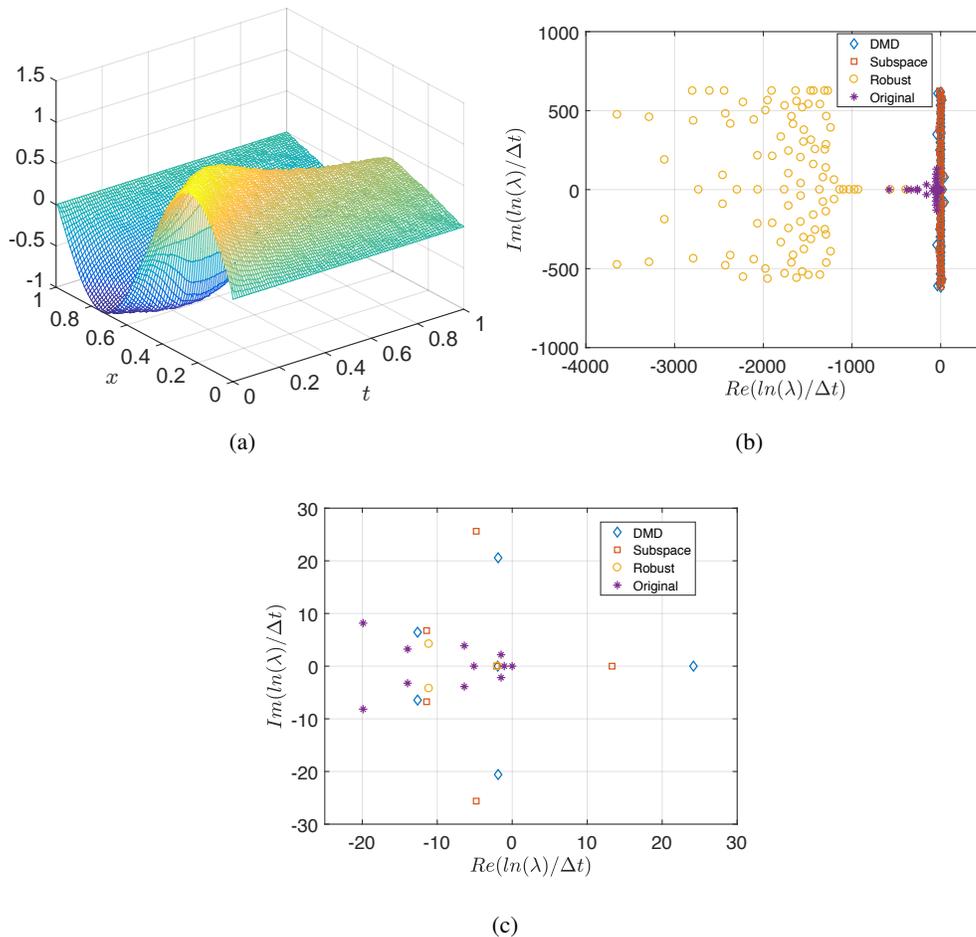
Figure 9.9 (a) Data generated by Stochastic Burger's equation with observation noise (b) Eigenvalues of the continuous time system.(c) Zoomed in region of dominant eigenvalues captured by the three methods.

Since the eigenvalues obtained with robust DMD algorithm shows a closer match, it is logical to believe that prediction using the robust DMD Koopman operator will be much more efficient compared

to either regular DMD or subspace DMD.
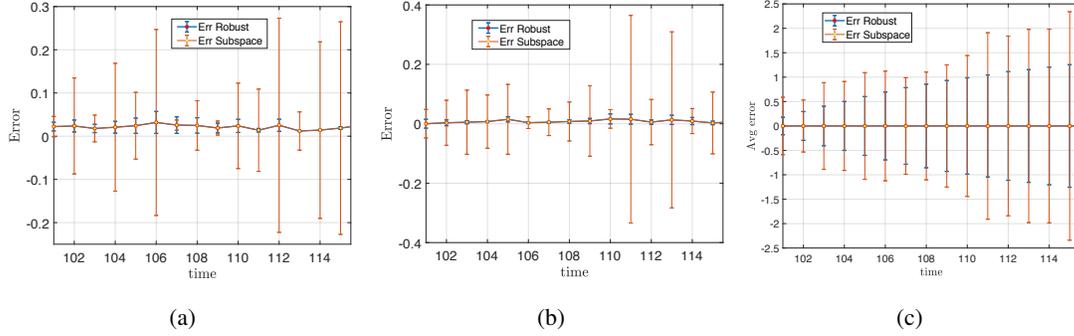
Figure 9.10 *(a) Error in prediction of z_2. (b) Error in prediction of z_{50}. (c) Average error in prediction.*

We used 100 time steps data as the training data and used the obtained operators to predict future 15 states. Since $\Delta x = 1 \times 10^{-2}$, the discretized system has 100 states, namely, $\begin{pmatrix} z_1 & z_2 & \cdots & z_{100} \end{pmatrix}$. In Fig. 9.10(a) and (b) we plot the errors in prediction of z_2 and z_{50} respectively, when using robust DMD and subspace DMD algorithms. It can be seen that the error in prediction with robust DMD is much smaller than the error obtained with subspace DMD. In Fig. 9.10(c) we plot the average error in all the states at each prediction time step. We do not plot the error using regular DMD because since the Koopman operator obtained using regular DMD has an eigenvalue with a large positive real part, it is highly unstable and the error diverges exponentially fast.

9.3   Causal Inference in Linear Dynamical System

In this section, we outline the procedure for computing the information-based causal inference from time series data for the case of linear dynamical system. We will employ the data-driven approximation of transfer operators discussed in the previous section and also exploit the fact that the analytical expression for information transfer in linear system are available. For ease of understanding, we discuss the procedure of information transfer computation for a two dimensional or two subspace case, as given by system Eq. (4.2). The general case will follow from the two subspace case.

Note that in all the subsequent discussion we assume that we have access to all the states of the system. The problem of computing the information transfer based on output measurements could be

more realistic problem and is a topic of our ongoing investigation. Let the time series data be given by

$$\mathcal{D} = \left[\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \cdots, \begin{pmatrix} x_{M-1} \\ y_{M-1} \end{pmatrix} \right] \tag{9.28}$$

Since the data is assumed to be generated from a linear dynamical system, we use linear dictionary functions i.e., $\psi_k(z) = z_k$. Furthermore, the number of dictionary functions are taken to be equal to the size of the system i.e., N. For the linear system case we use optimization formulation (9.19) i.e., EDMD with linear dictionary functions for the approximation of Koopman operator. With the linear choice of dictionary function and number of dictionary function equal to size of the system it is not difficult to show that the approximation of Koopman operator, $\mathbf{K}$, is the system A matrix itself. Let $\bar{A} = \mathbf{K} \in \mathbb{R}^{\mathbf{N} \times \mathbf{N}}$ be the estimated system dynamics obtained using optimization formulation (9.19). Under the assumption that the initial covariance matrix is $\bar{\Sigma}(0)$, the propagation of the covariance matrix under the estimated system dynamics $\bar{A}$ is given by

$$\bar{\Sigma}(t) = \bar{A}\bar{\Sigma}(t-1)\bar{A}^\top + \sigma^2 I \tag{9.29}$$

Both $\bar{A}$ and $\bar{\Sigma}$ can be decomposed according to Eqs. (4.3) and (4.4). The conditional entropy $H(y_{t+1}|y_t)$ for the non-freeze case is computed using the following formula (96; 97).

$$H(y_{t+1}|y_t) = \frac{1}{2} \log |\bar{A}_{yx}\bar{\Sigma}_y^S(t)\bar{A}_{yx}^\top + \left(\frac{\lambda}{3}\right)^2 I|. \tag{9.30}$$

where $|\cdot|$ is the determinant, λ is the bound on the process noise, $\bar{\Sigma}_y^S(t)$ is the Schur complement of y in the covariance matrix $\bar{\Sigma}(t)$. In computing the entropy, we assume that the noise is i.i.d. Gaussian with covariance $\Sigma = \text{diag}(\sigma^2, \cdots, \sigma^2)$ so that one can take the bound as $\lambda = 3\sigma$, to cover the essential support of the Gaussian distribution.

Computing the conditional entropy of y when x is frozen from the time series data obtained from the non-freeze dynamics is a challenge. To replicate the effect of x freeze dynamics we modify the original data set (9.28) as follows.

$$\mathcal{D}_{\not{x}} = \left[\left\{ \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \begin{pmatrix} x_0 \\ y_1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_1 \\ y_2 \end{pmatrix} \right\}, \cdots, \left\{ \begin{pmatrix} x_{M-1} \\ y_{M-1} \end{pmatrix}, \begin{pmatrix} x_{M-1} \\ y_M \end{pmatrix} \right\} \right] \tag{9.31}$$

If the original data set has M data points, then the modified data set has $(2M - 2)$ data points. The idea is to find the best mapping that propagate points of the form $[x_{t-1} \quad y_{t-1}]^\top$ to $[x_{t-1} \quad y_t]^\top$ (i.e., x freeze) for $t = 1, 2, \ldots, M$. The estimated dynamics $\bar{A}_{\not{y}}$, when x is frozen, is calculated using the optimization formulation (9.17) but this time applied to the data set (9.31). Once the frozen model is calculated, the entropy $H_{\not{y}}(y_{t+1}|y_t)$ is calculated using exactly the same procedure outline for $H(y_{t+1}|y_t)$ but this time applied to $\bar{A}_{\not{y}}$. Finally the information transfer from $x \to y$ is computed using the formula

$$T_{x \to y} = H(y_{t+1}|y_t) - H_{\not{y}}(y_{t+1}|y_t)$$

The algorithm for computing the information transfer for the linear system case can be summarized as follows:

Algorithm 1   Information Transfer: Linear System

---

1. From the original data set (9.28), compute the estimate of the system matrix $\bar{A}$ using the optimization formulation (9.19)

2. Assume $\bar{\Sigma}(0)$ and compute $\bar{\Sigma}(t)$ using Eq. (9.29). Determine $\bar{A}_{yx}$ and $\bar{\Sigma}_y^S$ to calculate the conditional entropy $H(y_{t+1}|y_t)$ using (9.30).

3. From the original data set (9.28) form the modified data set for the x freeze dynamics as given by Eq. (9.31).

4. Follow steps (1)-(2) to compute the conditional entropy $H_{\not{y}}(y_{t+1}|y_t)$.

5. Compute the transfer $T_{x \to y}$ as $T_{x \to y} = H(y_{t+1}|y_t) - H_{\not{y}}(y_{t+1}|y_t)$.

Example 40. *In the first example, we discuss the physical meaning of information transfer and demonstrate how it can be used to characterize influence in a dynamical system. Consider a mass-spring-damper system, as shown in Fig. 9.11.*

The equations of motion for the mass-spring system are

$$M\ddot{x}_1 + 2d\dot{x}_1 - d\dot{x}_2 + 2kx_1 - kx_2 = 0 \tag{9.32}$$

$$m\ddot{x}_1 + 2d\dot{x}_2 - d\dot{x}_1 + 2kx_2 - kx_1 = 0 \tag{9.33}$$
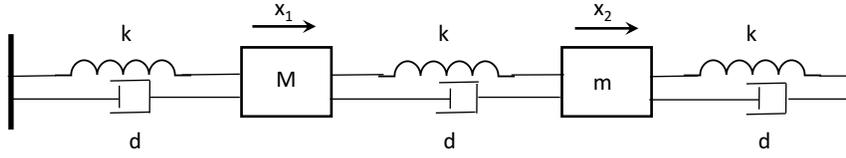
Figure 9.11   Mass-spring-damper system

where M, m are the masses, d is the damping coefficient and k is the spring constant. We assume that the damping coefficients of the dampers are equal and so are the spring constants of the springs. In state space form, the system can be represented as

$$\begin{pmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -2k/M & -2d/M & k/M & d/M \\ 0 & 0 & 0 & 1 \\ k/m & d/m & -2k/m & -2d/m \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}$$

where $z_1 = x_1$, $z_2 = \dot{x}_1$, $z_3 = x_2$ and $z_4 = \dot{x}_2$. For simulation purposes, we choose $M = 10$, $m = 1$, $k = 1$ and $d = 5$ with appropriate units. Since $M > m$, a perturbation (perturbed so that it has some non-zero initial velocity) in M will result in larger oscillations in the masses, compared to the case when m is perturbed by the same amount. Hence, we can conclude that M has a large influence on m, whereas, m has much smaller influence on M. In the language of information transfer between the states, this can be characterized by the information transfer from the position variable of one mass to the velocity variable of the other mass. From the analytical expression of information transfer, that is, equation (4.6), the information transfer from $z_1 \rightarrow z_4$ and $z_3 \rightarrow z_2$ are $T_{z_1 \rightarrow z_4} = 0.174$ and $T_{z_3 \rightarrow z_2} = 0.053$. This confirms the fact that M has a much larger influence on m, whereas m has negligible effect on M.

For calculating these transfers from data, the system was initialized at $\begin{bmatrix} 1 & .2 & 1 & .3 \end{bmatrix}^\top$ and data was collected for 10 time steps with sampling time $\delta t = 0.1$ seconds. An additive Gaussian noise of variance 0.1 was added to the system. With this, the information transfer values were calculated as $T_{z_1 \rightarrow z_4} = 0.1502$ and $T_{z_3 \rightarrow z_2} = 0.03$. Though the exact values do not match, but they are close

and more importantly, they do convey the fact that M has a much larger influence on m and m has negligible influence on M.

9.3.1 Topology Identification of Linear Networks

For linear systems, zero information transfer can be characterized by the entries of the system matrix and in particular, information transfer from a state z_i to z_j is zero if and only if A_{ji} is zero. Hence, a non-zero information transfer from z_i to z_j implies existence of a directed link from node z_i to z_j in the underlying linear network. We use this result and use information transfer as a tool for identification of topology of linear networks.

Example 41. *In this example the objective is to identify network topology from time series data. Consider a network dynamical system described by following difference equation*

$$\begin{pmatrix} z_{t+1}^1 \\ z_{t+1}^2 \\ z_{t+1}^3 \\ z_{t+1}^4 \\ z_{t+1}^5 \end{pmatrix} = 0.9 \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} z_t^1 \\ z_t^2 \\ z_t^3 \\ z_t^4 \\ z_t^5 \end{pmatrix} + \sigma \xi_t \tag{9.34}$$

The network topology corresponding to above system is shown in Fig. 9.12

We know that the information transfer from $T_{x_1 \to y}$ is zero if and only if $A_{yx_1} = 0$. Hence, information transfer can be used to inference presence or absence of a link i.e., network topology. In the above system z^i dynamics affect the dynamics of z^{i+1} for $i = 1, 2, 3, 4$ and z^1 is affected by z^3. Hence, $T_{z^i \to z^{i+1}}$ should be non-zero for $i = 1, 2, 3, 4$ and $T_{z^3 \to z^1}$ should also be non-zero. All the other transfers should be zero. The data was generated by choosing a single initial condition and propagating it for ten time steps with value of noise variance $\sigma = 2.1$.

The information transfer between the states is shown in table 9.1 and Fig. 9.13 and we find that there is non-zero information transfer from $z^i \to z^{i+1}$ for $i = 1, 2, 3, 4$ and $T_{z^4 \to z^1}$ is also non-zero. All the other transfers are very close to zero. For example, the transfer from z^5 to all the other states is of the order of 10^{-4} and hence we conclude that z^5 is not affecting any other state. Hence, we find
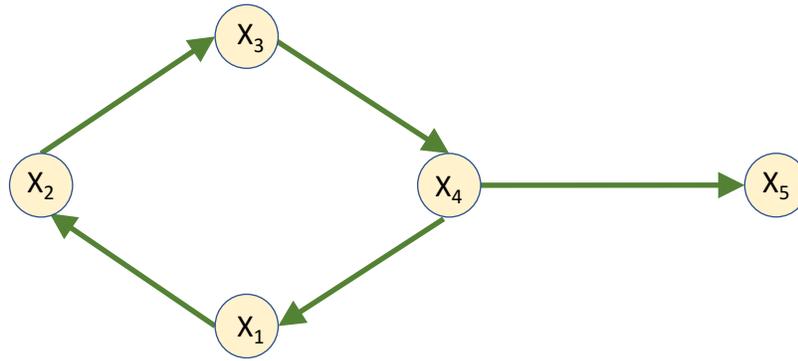
Figure 9.12   Network corresponding to the dynamical system given in (9.34).
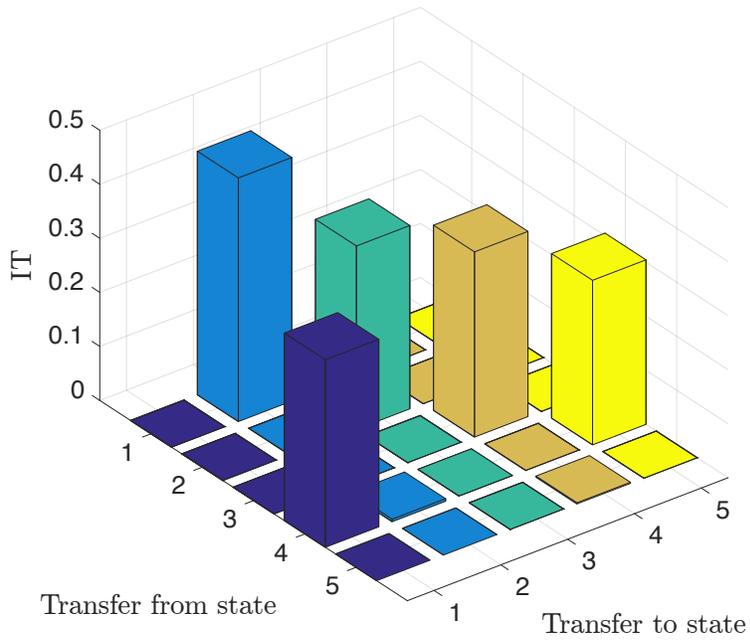
Figure 9.13   Information transfer between the states.

Table 9.1   Information Transfer between the States

I.T.	Value
$T_{z_1 \to z_2}$	0.45
$T_{z_2 \to z_3}$	0.34
$T_{z_3 \to z_4}$	0.34
$T_{z_4 \to z_5}$	0.31
$T_{z_4 \to z_1}$	0.35
Other transfers	$\sim 10^{-4}$

that our information transfer measure recovers the correct causal structure or the network structure of the dynamical system from time series data.
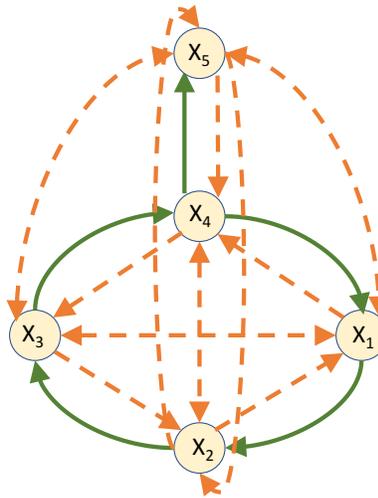
Figure 9.14   Network identified by Granger causality

Further, to compare with an existing measure of causality, we used Granger causality test (8),(5) on the same data set. Granger causality test is one of the most commonly used methods for causality detection and the intuition behind the definition is the following. A variable X Granger causes another variable Y if the prediction of Y based on its own past and past of X is better than the prediction of Y based on its own past alone. For self containment of this work, we discuss the formulation of Granger causality briefly. Suppose X_t, Y_t and Z_t are three jointly distributed stationary multivariate stochastic processes. Consider the regression models

$$X_t \;=\; \alpha_t + (X_{t-1}^{(p)} \oplus Z_{t-1}^{(r)}) \cdot A + \epsilon_t \tag{9.35}$$

$$X_t \;=\; \alpha'_t + (X^{(p)}_{t-1} \oplus Y^{(q)}_{t-1} \oplus Z^{(r)}_{t-1}) \cdot A' + \epsilon'_t \tag{9.36}$$

where A and A' are the regression coefficients, α and α' are constant terms, ϵ and ϵ' are residuals, and the predictee variable X is regressed first on the previous p lags of itself plus r lags of the conditioning variable Z and second, in addition, on q lags of the predictor variable Y. Granger causality of Y to X, given Z, is a measure of of the extent to which the inclusion of Y in the model (9.36) reduces the prediction error of the first model (9.35) and is defined as

$$G_{Y \to X|Z} = \ln \frac{\mathrm{var}(\epsilon_t)}{\mathrm{var}(\epsilon'_t)} \tag{9.37}$$

where $\mathrm{var}(\cdot)$ is the variance.
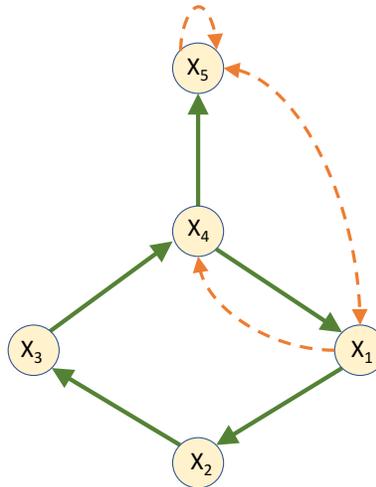
Figure 9.15   Network obtained using Sparse DMD

As can be seen from Fig. 9.14, Granger causality identifies both direct and indirect causal influence and it fails to differentiate between direct and indirect influence. The direct links are shown in green and the indirect links, which are also identified by Granger causality are indicated by dotted orange lines. Hence, it is not possible to infer the correct causal structure from Granger causality test. However, our measure of information transfer can differentiate between direct and indirect influence (96) and information transfer computed in this work gives only the direct influence. Hence the proposed measure captures the true causal structure.

At this stage one might wonder as to the need of computing information transfer for the purpose of identifying network topology. If all we are interested in determining the presence of absence of links between nodes, one can do that by simply estimating the system dynamics matrix $\bar{A}$ using optimization formulation in Eq. (9.19) from time-series data. In that case, one may wonder is there a need to estimate $\bar{A}_{\notin}$ for the freeze system? In particular, if $\bar{A}_{ij} = 0 (\neq 0)$ then it implies absence (presence) of directed link from node j to node i. To verify this claim in Fig. 9.15, we compare the results for the network topology identification obtained using our information transfer based method and one based on estimated system matrix $\bar{A}$. We find that the information transfer measure can regenerate the correct topology of the network, whereas, sparse DMD algorithm identifies some links which are not there in the original network. The spurious links are marked by dotted orange lines in Fig. 9.15.

Example 42. *Small world networks are ubiquitous in nature and in this example, we look at a small world network with 20 nodes and analyze how information transfer measure performs to recover the causal structure of the small world network. In the previous example, we constructed a network from a given dynamical system. In this example, we start with a network and construct a dynamical system from the network and identify the connections of the network using the information transfer measure. In particular, given a network of n nodes, we construct a n-dimensional discrete time linear dynamical system such that, if there is a directed edge from node i to node j, then $A_{ji} \neq 0$. Thus we obtain the system matrix of the dynamical system and we make the system stable by scaling the A matrix with an appropriate constant.*

In Fig. 9.16(a) we show the small world network of 20 nodes which is used to generate the data. In this case, we corrupt the data with i.i.d. Gaussian noise of variance 0.1. Hence, $\lambda = 0.3$ in the optimization problem (9.17). Fig. 9.16(b) shows the reconstructed network using information transfer method. We find that in this case information transfer identifies three extra links, which are not there in the original system. These links are marked in red in Fig. 9.16(b). In this case, information transfer does not identify the exact causal structure and there is some error, but it performs pretty well.

In Fig. 9.17, we show the percentage error as the number of nodes in the system is increased. In all these cases, we did not corrupt the data with noise. In these examples, we find that even as the number of nodes is increased, the percentage error remains relatively small, and hence we conclude
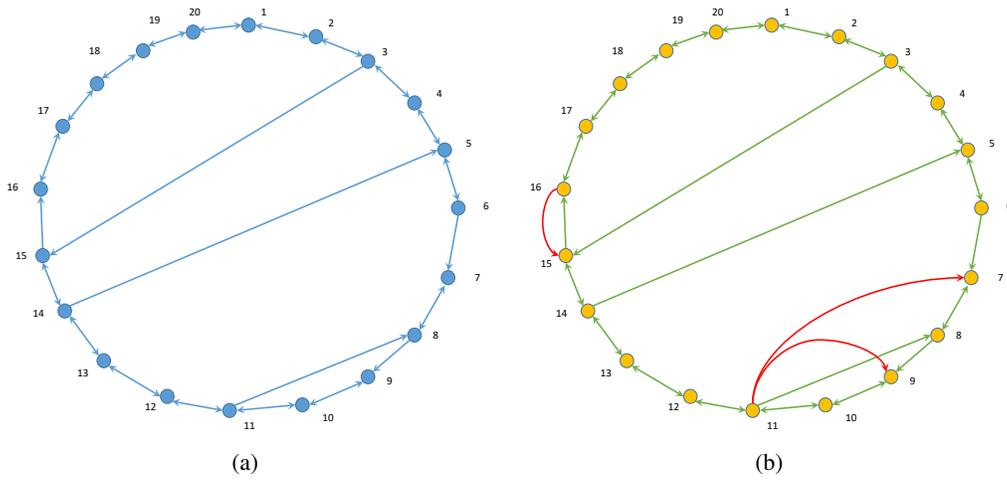
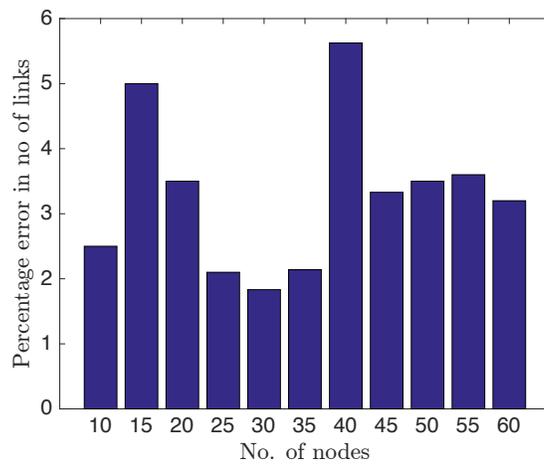*Figure 9.16   (a) Small world netowrk of 20 nodes. (b) Reconstructed network.*

Figure 9.17   Percentage error in number of links v/s number of nodes.

that the information transfer measure does a pretty good job in identifying the causal structure in a dynamical system. In these simulations, we considered small world networks of different number of nodes and the number of links is nearly four times the number of nodes.

9.4 Computation of the Information Transfer for Nonlinear Systems

One of the main challenge in computing the information transfer from time series data for nonlinear system is to propagation of probability density function under nonlinear flow field and the computation of conditional entropy term. Unlike linear system where linearity and Gaussian property of the probability density function was exploited for the data-driven computation of information transfer, the same does not applies for nonlinear system. We make use of Naturally Structured Dynamic Mode Decomposition algorithm (NSDMD) for the approximation of transfer P-F operator and for propagation of probability density function. The positivity and the Markov property of the NS-DMD algorithm is exploited for the propagation of probability density function and for computation of conditional entropy term.

Again we will outline the computation procedure for the two subspace case the general case (more than two subspace case) will follows from this procedure. The information transfer formula for the two subspace case can be simplified as follows. We rewrite Eq. (3.3) as

$$T_{x \to y} = H(\rho(y(t+1)|y(t))) - H(\rho_{\not x}(y(t+1)|y(t)))$$
$$= H(\rho(y(t+1), y(t))) - H(\rho_{\not x}(y(t+1), y(t)))$$

In writing the above equality we have used the fact that $H(X|Y) = H(X,Y) - H(Y)$. Furthermore, since x is held frozen from time t to time $t+1$, we have

$$H(\rho(y(t))) = H(\rho_{\not x}(y(t)).$$

We next outline the procedure for computing the joint entropy term $H(\rho(y(t+1), y(t)))$ using the finite-dimensional approximation of the P-F matrix $\mathbf{P}$ obtained using NSDMD optimization formulation outlined in (9.15). Note that in the construction of the P-F matrix $\mathbf{P}$ we use the original data set (9.28). Once we outline the procedure for computing the joint entropy term for the non freeze case, $H(\rho(y(t+1), y(t)))$, the computation for the entropy term for the freeze case, $H(\rho_{\not x}(y(t+1), y(t)))$,

will follow along similar lines. The only difference being the the P-F matrix for the freeze case, denoted by $\mathbf{P}_{\not\not}$, will be computed using the modified data set obtained to replicate the freeze case i.e., data-set (9.31).

For computing $H(\rho(y(t+1), y(t)))$, we first consider finite approximation of $\rho(z(t+1), z(t))$ as discrete probability measure. Towards this goal we consider finite partition of the state space Z as

$$\mathcal{Z} = \{D_1, \ldots, D_K\}, \quad \mathcal{D} = \cup_{k=1}^{K} D_k.$$

such that $D_i \cap D_j = \emptyset$. Similarly, let

$$\mathcal{Z}^x = \{D_1^x, \ldots, D_K^x\}, \quad \mathcal{Z}^y = \{D_1^y, \ldots, D_K^y\}$$

where $\mathcal{Z}^x$ and $\mathcal{Z}^y$ are the projection of the partition, $\mathcal{Z}$, along the x and y coordinates respectively. Let $\mathcal{D}^x = \cup_{k=1}^{K} D_k^x$ and $\mathcal{D}^y = \cup_{k=1}^{K} D_k^y$. We have $\rho(z(t+1), z(t)) = \rho(z(t))\rho(z(t+1)|z(t))$, Let

$$[p_z]_{ij}^t := \text{Prob}(z_{t+1} \in D_j | z_t \in D_i)$$

$$[p_z]_i^t := \text{Prob}(z_t \in D_i)$$

$$\text{Hence,} \quad \text{Prob}(z_{t+1} \in D_j, z_t \in D_i) = [p_z]_i^t [p_z]_{ij}^t$$

Similarly, we can define $[p_x]_i^t$, $[p_y]_i^t$, $[p_x]_{ij}^t$, and $[p_y]_{ij}^t$ as follows:

$$[p_x]_i^t := \text{Prob}(x_t \in D_i^x), \quad [p_y]_i^t := \text{Prob}(y_t \in D_i^y)$$

To compute the above defined quantities, we make use of finite dimensional approximation of P-F operator, $\mathbf{P}$. Note that in the finite dimensional approximation of the P-F matrix using NSDMD algorithm we have assumed that the dictionary functions are positive i.e., $\psi_i(z) \geq 0$ for $i = 1, \ldots K$. Furthermore, we also assume that the dictionary functions are density functions i.e.,

$$\int_Z \psi_i(z)dz = 1, \quad i = 1, \ldots, K.$$

Let $w_t = (w_t^1, \ldots, w_t^K) \in \mathbb{R}^K$ be a probability row vector. Density function, $\rho(z(t))$ can be constructed using this probability vector and the dictionary functions $\mathbf{\Psi}(z)$ as $\rho(z(t)) = w_t \mathbf{\Psi}^\top(z)$. This density function can be propagated using finite dimensional $\mathbf{P}$ as follows.

$$\rho(z(t+1)) = w_t \mathbf{P}^\top \mathbf{\Psi}^\top(x) = w_{t+1} \mathbf{\Psi}^\top(x).$$

Hence we have

$$[p_z]_i^t = \int_{D_i} w_t \mathbf{\Psi}^\top(z)dz = w_t \int_{D_i} \mathbf{\Psi}^\top(z)dz = w_t \Theta_i^\top$$

where $\Theta_i = \int_{D_i} \mathbf{\Psi}(z)dz$.

Let $\lambda_i = \{i_1, i_2, \cdots, i_L\} \subseteq \{1, 2, \cdots, K\}$ such that $D_{i_k}^y \cap D_i^y \neq \phi$, where ϕ is the empty set. Then

$$[p_y]_i^t = \int_{D_{\lambda_i}} w_t \mathbf{\Psi}^\top(z)dz = w_t \int_{D_{\lambda_i}} \mathbf{\Psi}^\top(z)dz.$$

Now let $\bar{w}$ be the probability vector such that

$$\bar{w} \int_{D_i} \mathbf{\Psi}^\top(z)dz = 1.$$

i.e., density function of the form $\bar{w}\mathbf{\Psi}^\top(z)$ correspond to the case where the entire distribution is concentrated on set D_i. With the above definition of $\bar{w}$ we have

$$[p_z]_{ij}^t = \int_{D_j} \bar{w}\mathbf{P}^\top \mathbf{\Psi} dz = \bar{w}\mathbf{P}^\top \int_{D_j} \mathbf{\Psi}(z)dz = \bar{w}\mathbf{P}^\top \Theta_j^\top$$

Hence, we have

$$\begin{aligned}
&\mathrm{Prob}(z_{t+1} \in D_j, z_t \in D_i)] \\
= \ & \left[w_t \int_{D_i} \mathbf{\Psi}^\top(z)dz\right] \left[\bar{w}\mathbf{P}^\top \int_{D_j} \mathbf{\Psi}(z)dz\right] \\
= \ & w_t \Theta_i^\top \bar{w}\mathbf{P}^\top \Theta_j^\top = \Gamma_{ij}
\end{aligned} \tag{9.38}$$

As defined earlier, the set $\mathcal{D} = \{\psi_1, \cdots, \psi_K\}$ are the dictionary functions for observables on the space Z. Note that $\rho(z(t+1), z(t))$ is defined on the product space $Z \times Z$ and hence we consider the set $\Phi = \mathcal{D} \times \mathcal{D} = \{\varphi_{11}, \varphi_{12}, \cdots, \varphi_{KK}\}$ as the set of dictionary functions on the product space, where $\varphi_{ij} = \psi_i \psi_j$ and let $D_i \times D_j := D_{i,j}$. Hence, we have

$$\rho(z(t+1), z(t)) = \sum_{i,j=1}^{K} \Gamma_{ij}\varphi_{ij}(z, w). \tag{9.39}$$

Let

$$\lambda_{i,j} = \{(p,k)|D_{i,j}^y \cap D_{p,k}^y \neq \phi; p, k = 1, 2, \cdots, K\}$$

Hence, by similar arguments for finding the marginal probability,

$$\text{Prob}(y(t+1) \in D_j, y(t) \in D_i)$$
$$= \int_{D_{\lambda_{i,j}}} \sum_{i,j=1}^{K} \Gamma_{ij} \varphi_{ij}(z, w) dz dw \tag{9.40}$$

Hence, (9.40) gives the probability distribution of $(y(t+1), y(t))$ and using the entropy formula for a discrete probability distribution $(H(P) = -\sum_i P_i \log P_i)$, we get the entropy of $(y(t+1), y(t))$. Similarly, using the same above procedure for the modified data set (9.31), we can compute the entropy of $(y(t+1), y(t))$, when x is held frozen and computing the difference $H(y(t+1), y(t)) - H_{\not x}(y(t+1), y(t))$, we get the information transfer from x to y.

Algorithm 2 Algorithm for finding the information transfer from time series data

1. Compute the Koopman operator from the time series data using the method of Naturally Structured Dynamic Mode Decomposition.

2. Compute the joint probability of $(z(t+1), z(t))$ using (9.38).

3. Compute the marginal probability of $(y(t+1), y(t))$ from the joint density from (9.40).

4. Compute the entropy $H(y(t+1), y(t))$.

5. Form the modified data set from the given time series data from (9.31).

6. Repeat steps (1)-(5) for the modified data set to get the entropy $H_{\not x}(y(t+1), y(t))$.

7. Compute $H(y(t+1), y(t)) - H_{\not x}(y(t+1), y(t))$ to get the information transfer from x to y.

9.4.1 Examples and Simulations

Example 43. Two State Non-linear System.

Next we consider a non-linear example. Consider the system

$$x_{t+1} = 2x_t(1 - x_t) + 2y_t; \quad y_{t+1} = .8y_t$$

The system was evolved for 300 time steps, starting from $\begin{pmatrix} 0.9 & 0.9 \end{pmatrix}^{\top}$.

From the system equations, we see that the y dynamics is not affected by x, whereas, x dynamics is affected by y. So there should be non-zero flow of information from y to x and there should be zero information flow from x to y.

For this example too, we considered Gaussian radial basis functions, with $\sigma = 0.01$, for computation of the Koopman operators and with our algorithm we found $T_{x \to y} = -0.03$ and $T_{y \to x} = 1.63$. So, in this case also our information transfer measure identifies that y affects x dynamics and x does not influence y. Hence we have identified the correct causal structure.

Granger causality for this example identified a statistical dependence of x and y and inferred that both x and y cause each other, whereas, in reality the y dynamics is never affected by x dynamics and hence the influence of x on y should be zero. However, Granger causality fails to identify this, while our information transfer measure does capture the zero influence.

Example 44. Henon Map. *In the second example, we consider the Henon map. It is one of the most studied dynamical systems which exhibit chaotic behaviour. The dynamical equations of the Henon map are*

$$x_{n+1} = 1 - ax_n^2 + y_n + \gamma \xi_x$$
$$y_{n+1} = bx_n + \gamma \xi_y$$

The classical values of the parameters are $a = 1.4$ and $b = 0.3$. We add small process noise in the system with $\gamma = 0.01$. The support of the the first two Koopman eigenfunctions, computed using the NSDMD algorithm, is shown in Fig. 9.18.
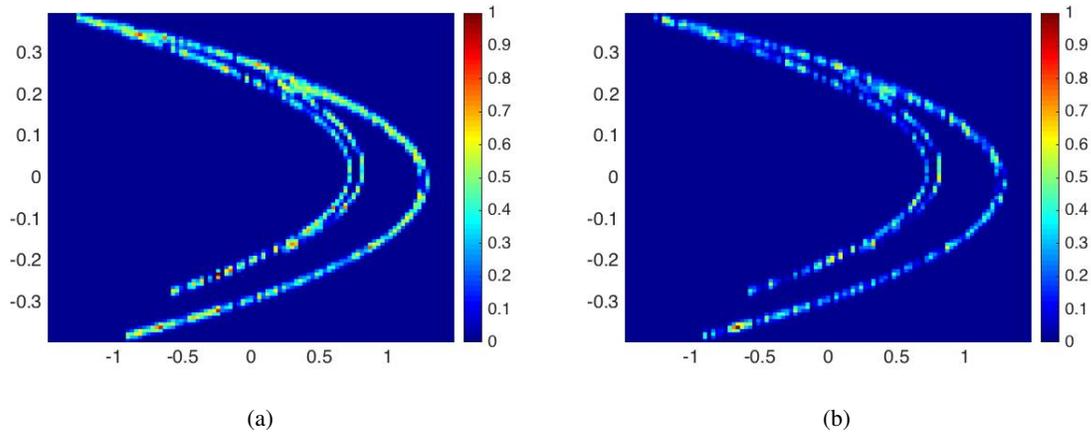
Figure 9.18 (a) Support of the eigenfunction corresponding to the largest eigenvalue of the Koopman operator. (b) Support of the eigenfunction corresponding to the second largest eigenvalue of the Koopman operator.

For computing the information transfers, data was collected for 1000 time steps and information transfer between the states was calculated using the algorithm described in algorithm 2. We chose 200 Gaussian radial basis functions, with $\sigma = 0.01$ as the dictionary functions and this gave $T_{x \to y} = 0.0982$ and $T_{y \to x} = -0.7246$. It is to be noted that one of the transfers is negative. We suspect that the negative value of the transfer is due to the fact that one Lyapunov exponent of the Henon map is negative, while the other is positive. Another interesting observation is the fact that the information transfer from x to y is nearly b^2. This is observed to be true in linear systems which have the following structure

$$
\begin{aligned}
x_{n+1} &= a_x x_n + a_{xy} y_n + \sigma \xi_x \\
y_{n+1} &= a_{yx} x_n + \sigma \xi_y
\end{aligned}
$$

In particular, it was observed that with small $\sigma \approx 0.01$, the information transfer from x to y is nearly equal to a_{yx}^2. Since the x dynamics affect the y dynamics linearly in the case of a Henon map, the same result hold true here as well.

CHAPTER 10. DATA DRIVEN CAUSAL INFERENCE IN STOCK MARKET

10.1 Introduction

Economists and common man alike are interested in stock market and stock prices of different companies. Economists try to explain and predict the behaviour of stock prices, whereas people who invest in the stock market are equally interested, as they have their property at stake. An organized and managed stock market also encourage investment opportunities (98; 99). Moreover, stock market plays an important role in the allocation of capital to corporate sector that in turn stimulate real economic activity. As such, studies related to cause and effect in stock markets are of utmost practical importance.

Identification of cause-effect relation has been a fundamental question in science since long. Causality and influence characterization is important in many different branches of research like biology, neuroscience, economics, social network analysis etc. Usually, ideas from statistics and information theory are used for characterizing causality and influence. For example, in (1) the authors use information based metric to characterize the most influential nodes in social networks. In neuroscience, concepts of information theory are used to understand how information flows in different parts of the brain (2) and identifying influence in gene regulatory networks (3; 4). In economic and financial networks, different causality tests can be used to infer causal interactions from the time series data. Among these Granger causality (6; 7; 8) is one of the most commonly used definitions of causality and Granger causality and different variants of it have been used extensively for characterizing influence in stock market (100; 101; 102).

In this chapter, we use the notion of information transfer in a dynamical system to analyze the causal structure among the stock prices of 84 different companies in the US market. In particular, we assume the stock market to be a discrete time linear system and use data driven approach to compute

the information transfer between the different companies and based on the information transfer values we characterize the causal structure and influence among the companies. Further, we use the information transfer measure to define what we call influence distance and demonstrate how it can be used for clustering of the companies. We also show how the information transfer measure can be used for prediction of stock market crashes and we demonstrate this by predicting the Great Recession of 2008 and the stock market crash of November 2008.

10.2 Influence in Stock Market

In this section, we discuss how the information transfer measure can be used for analysis of stock market prices. U.S. closing stock prices of 84 different companies over a period of 600 days was collected for analysis. In particular, closing stock prices from 3rd January 2007 to 9th June 2009 was used for analysis. Data from this period was chosen so that it includes the data from the Great Depression of 2007-08 and also data from financial collapse in November 2008.

10.2.1 Influence between 2 companies

Information transfer from a state x to any other state y quantifies how the x dynamics affect y dynamics, as the system evolves in time. For example, if the absolute value of information transfer from x to y is large, it means that x dynamics affect (influence) y a lot. This is because, large information transfer means large transfer of entropy and so, as x evolves, it makes y fluctuate and hence y changes. Hence, if $|T_{x \rightarrow y}|$ is large, change in value of x will result in change in y. On the other hand, if $T_{y \rightarrow x} \approx 0$, then variation in y may not change x.

For example, we look at the information transfer between Internation Business Machines (IBM) and Intel Inc. (Intel). The first 300 days data (closing stock price) was used for computing the information transfer and the information transfers were computed as $T_{IBM \rightarrow INTC} = 0.02$ and $T_{INTC \rightarrow IBM} = .2$. Hence, Intel influences the stock price of IBM, but IBM does not affect the stock price of Intel. This can be seen from Fig. 10.1. For example, around the 80 day mark, there is a significant change in the stock price of IBM, but the price of stocks of Intel remains fairly constant. On the other hand, around 250 days mark, there is a change in the stock price of Intel and we find that
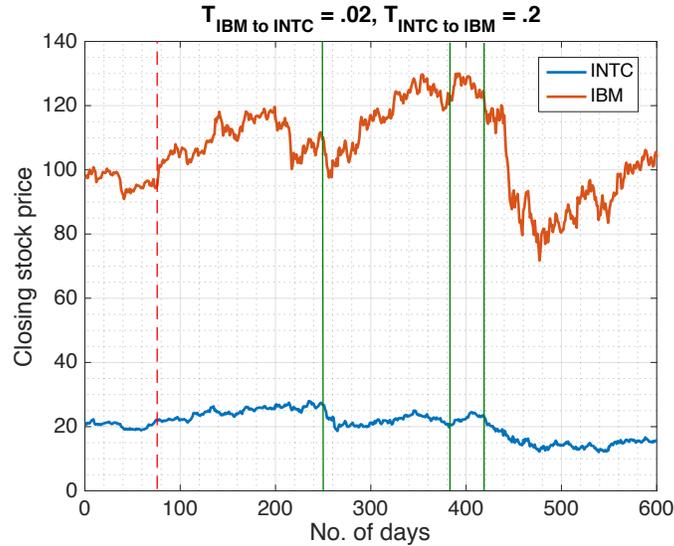
Figure 10.1 Stock prices of IBM and Intel (INTC).

stock price of IBM also changes. In fact, this observation is true around the 380 days mark and 420 days mark. Note that, for computing the information transfer, we had used the first 300 days stock price data. Hence, one can conclude that the computed information transfer can not only be used to verify the past data, but also to predict changes in stock prices in the future.

A similar conclusion can be deduced for Coca-Cola (KO) and Pepsi(PEP). The stock price data for the two companies is shown in Fig. 10.2. In this case also we took the first 300 days data as the training data and the information transfers were computed as $T_{KO \to PEP} = 0.23$ and $T_{PEP \to KO} = 0.03$. Hence, conclusions similar to the IBM and Intel analysis can be inferred. For example, around 440 days mark there is a sudden change in stock price of Coca-Cola and we see that there is a large drop in stock price of Pepsi as well. Whereas, around 500 days mark there is a change in stock price of Pepsi, but the stock price of Coca-Cola is not changing much.

As a third example, we look at stock prices of AT & T (T) and Verizon (VEZ). Using the first 300 days data as training data, the information transfer in both directions were found to be nearly 0.2. Hence, a change of stock price in each should reflect in a change in stock price of the other and the trend of the stock prices should be similar for both the companies. In fact, this is true, as can be seen in Fig. 10.3, where we plot the stock prices of both the companies.
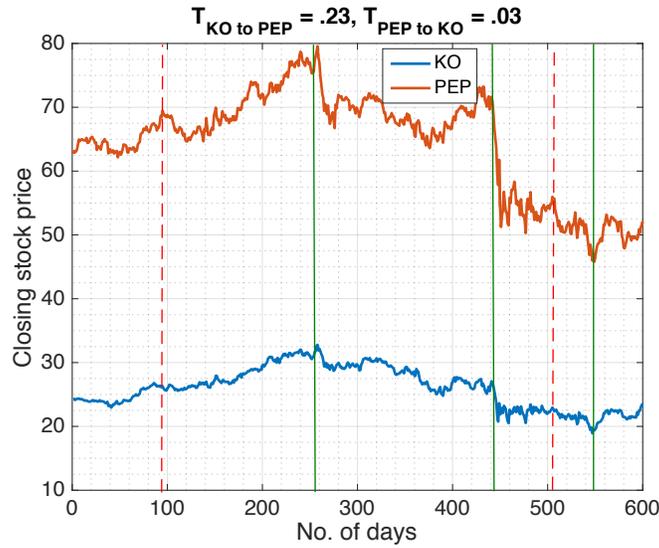
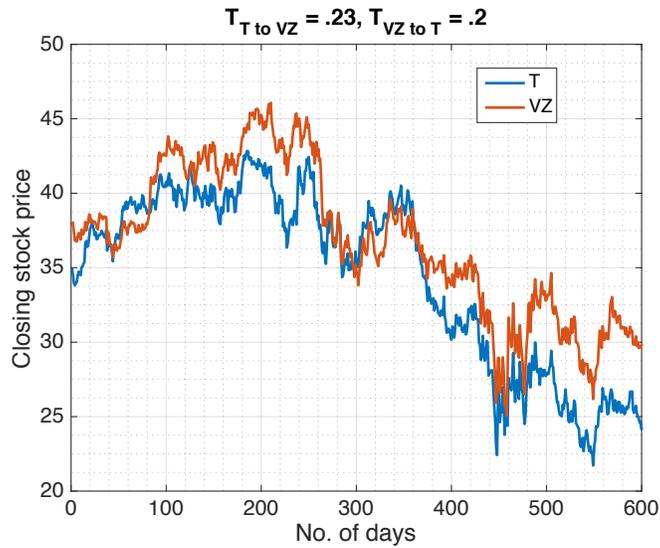Figure 10.2   Stock prices of Coca-cola (KO) and Pepsi (PEP).

Figure 10.3   Stock prices of At & T (T) and Verizon (VEZ).

10.2.2 Clustering of companies

Clustering of companies is important not only for investors, but also for creditors, stock holders etc. It is also an important tool for studying stock market volatility (103). As such a meaningful clustering of various companies is important. In this subsection, we present spectral and hierarchical clustering, based on the information transfer measure. Based on the information transfer $T_{x \to y}$ from x to y, we define the information distance or influence distance as follows :

Definition 45 (Influence Distance)**.** *If the information transfer from x to y is $T_{x \to y}$, then the information distance or influence distance from x to y is*

$$D_{x \to y} = \exp(-|T_{x \to y}|) \tag{10.1}$$

Note that the above measure is a metric which is non-symmetric and does not satisfy the triangle inequality. However, this kind of distance functions, which are not symmetric and does not satisfy triangle inequality is commonly used in information geometry (104). The rationale behind defining the distance in a manner as in (10.1) is the fact that if x has a large influence on y, then y is *close* to x.

With the distance defined as in (10.1), we look at the spectral clustering of the different banks, financial companies and computer companies (Fig. 10.4). We find that the major computer hardware companies like Microsoft (MSFT), Apple (AAPL) and Intel (INTC) are in the same cluster. Similarly, major financial companies like American Express (AXP), Capital One Financial (COF) and Schlumberger (SLB) belong to the same cluster. These basic findings indicate that the influence distance does make physical sense, and in the future, we hope to build on these findings. Moreover, a kind of crude verification of the clustering can be done by looking at the correlation between the stock prices of the companies in different clusters. For example, J.P. Morgan Chase (JPM) and Morgan Stanley (MS) belong to the same cluster and the correlation coefficient between the two, with the given data is 0.8521, whereas the correlation coefficient between JPM and Schlumberger Ltd. (SLB) is 0.5842 and they belong to different clusters.

In Fig. 10.5 we look at the hierarchical clustering of the same companies. This also clubs similar companies together. For example, Apple (AAPL) and Hewlett-Packard (HPQ) are in the same cluster,
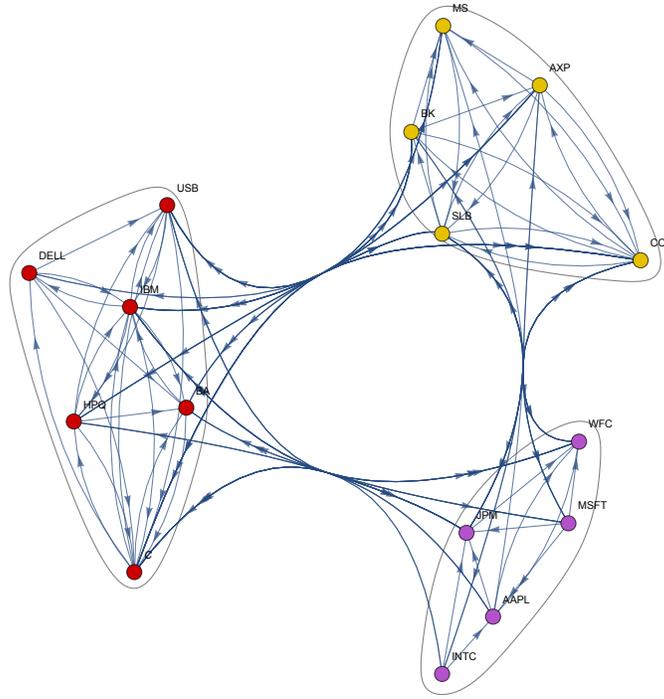
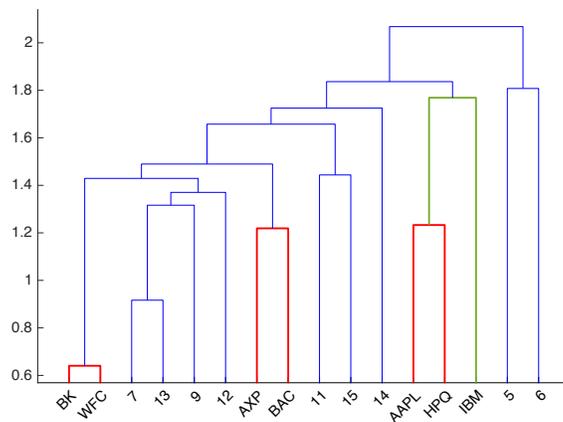Figure 10.4   Spectral clustering of different banks and computer companies.

Figure 10.5   Hierarchical clustering of different financial and computer companies.

which is superseded by International Business Machines (IBM). This is shown in Fig. 10.5 by red and green lines.

10.2.3 Stability of stock market

As mentioned earlier, in this work we have used stock price data of 600 days from January 2007 through June 2009. For stability analysis, the data was divided into 19 different windows, each containing data for 30 days. For stability analysis, we concentrate on the stock prices of 10 banks and financial companies and the information transfer between the companies was calculated for each of the 19 windows.
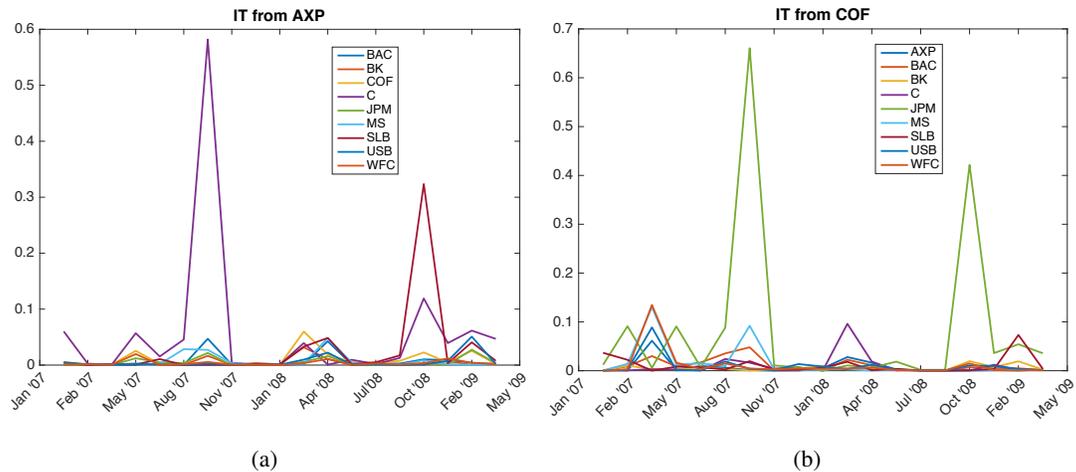
Figure 10.6   Information transfer from (a) American Express (AXP) and (b) Capital One Financial (COF).

The information transfer from American Express (AXP) and Capital One Financial (COF) is shown in Fig. 10.6(a) and (b) respectively. From Fig. 10.6(a) and (b), we find that there are two distinct peaks, one around September 2007 and another around September 2008. We know that in 2008 there was a severe recession, which is known as Great Recession of 2008 (105; 106). It is generally accepted that the recession started around late November and early December of 2007 (107). The large peak in information transfer obtained around September of 2007 predicts that instability is creeping into the system and one has enough time to take precautionary measures. Similarly, there was another crash around November 2008 (108). We find that information transfer can predict this

crash as well and that too in advance. However, this crash was not as severe as the Great Recession and this is reflected in the fact that the peak information transfer around September 2008 is smaller than the peak around September 2007. Hence, from the simulations, it is evident that information transfer can predict crashes in the stock market and hence can be used to analyze and predict stock prices and take corrective actions so that one does not suffer losses.

CHAPTER 11.   CONCLUSION AND FUTURE WORK

11.1   Conclusions

In this work, we studied the information flow between the states of a discrete-time dynamical system. The motivation for the work was to provide a definition of causality and influence in a dynamical system setting, so that the proposed measure can capture the intuitions of causality, like zero transfer and transfer asymmetry. To this end, we define what we call one-step information transfer and have shown that this definition does capture the intuitions of causality in a dynamical system. The one-step information transfer was generalized to define n-step information transfer and this allows complete distinction between direct and indirect transfer. Further, the definition of information transfer was generalized to define information transfer between the various signals in a control dynamical system. The definition of information transfer from state to output has physical meaning and we demonstrate this on power networks, where we use the information transfer measure to identify the generators and states which are responsible for instability. The same conclusions can also be reached from state to state transfer, where we connected information transfer with instability and this further allowed us to predict phase transitions in complex dynamical systems. We also provide two algorithms to compute the information transfer measure from time series data and use it for topology identification of linear networks and also characterized influence and also showed how the developed framework could be used to predict crashes in the stock market.

11.2   Future Work

The concepts presented here are only the starting point and can be extended in many different directions, both theoretically and from an application point of view. One of the main direction is

the connection between information transfer and thermodynamic concepts like entropy, free energy and temperature. In particular, can the laws of thermodynamics be stated/understood in terms of information transfer? Moreover, phase transition is a very important and relevant problem both in classical physics and quantum physics and a standard model for phase transitions is the double-well potential. The results on information transfer and phase transition show enough potential which can be explored in future. This connection not only has theoretical importance, but also practical implications in different fields like solid mechanics, chemistry, solid state physics etc. From the control theory perspective, information transfer can be used to design control and also design filters. In particular, one can show that in a Kalman filter the information transfer from the system to the estimator is a minimum. This shows a plausible direction in the future. Information transfer can be used to characterize influence and from an application point of view, this can be used in different real life problems like analysis of social networks, gene regulatory networks etc.

BIBLIOGRAPHY

[1] G. V. Steeg and A. Galstyan, "Information transfer in social media," in *Proceedings of the 21st international conference on World Wide Web*, New York, NY, 2012, pp. 509–518.

[2] O. Sporns, *The networks of the brain*. MIT Press, 2010.

[3] D. J. S. A. Rao, A. O. Hero and J. D. Engel, "Motif discovery in tissue-specific regulatory sequences using directed information." in *EURASIP J. on Bioinformatics and Systems Biology*, 2007.

[4] ——, "Inference of biologically relevant gene influence networks using the directed information criterion." in *In proc. ICASSP*, Toulouse, France, 2006.

[5] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[6] C.W.J. Granger, "Investigating causal relations by econometrics models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.

[7] C. A. Sims, "Money, income and causality," *American Economic Review*, vol. 62, pp. 540–552, 1972.

[8] C. W. J. Granger, "Testing for causality," *Journal of Economic Dynamics and Control*, vol. 2, pp. 329–352, 1980.

[9] T. Schreiber, "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461–464, July, 2000.

154

[10] X. S. Liang and R. Kleeman, "Information transfer between dynamical system components," *Physical Review Letters*, vol. 95, p. 244101, 2005.

[11] J. L. Massey, "Causality, feedback and directed information." in *Proc. Intl. Symp. on Info. th. and its Applications*, Waikiki, Hawai, USA, 1990.

[12] G. Kramer, "Directed information for channels with feedback," in *PhD Thesis*, Swiss Federal Institute of Technology Zurich, 1998.

[13] J. Sun and E. M. Bollt, "Causation entropy identifies indirect influences, dominance of neighbors and anticipatory couplings," *Physica D: Nonlinear Phenomena*, vol. 267, pp. 49–57, 2014.

[14] A. J. Majda and J. Harlim, "Information transfer between subspace of complex dynamical systems," *PNAS*, vol. 104, no. 23, pp. 9558–9563, 2007.

[15] X. San Liang and R. Kleeman, "A rigorous formalism of information transfer between dynamical system components. i. discrete mapping," *Physica D: Nonlinear Phenomena*, vol. 231, no. 1, pp. 1–9, 2007.

[16] ——, "A rigorous formalism of information transfer between dynamical system components. ii. continuous flow," *Physica D: Nonlinear Phenomena*, vol. 227, no. 2, pp. 173–182, 2007.

[17] X. S. Liang, "Local predictability and information flow in complex dynamical systems," *Physica D*, vol. 248, pp. 1–15, 2013.

[18] ——, "Information flow within stochastic dynamical systems," *Phys Rev E*, vol. 78:031113, 2008.

[19] H. Touchette and S. Lloyd, "Information-theoretic approach to the study of control systems," *Physica A: Statistical Mechanics and its Applications*, vol. 331, no. 1, pp. 140–172, 2004.

[20] Y. Tomita, S. Ohmatsu, and T. Soeda, "An application of the information theory to estimation problems," *Information and Control*, vol. 32, no. 2, pp. 101–111, 1976.

[21] S. K. Mitter and N. J. Newton, "Information and entropy flow in the kalman-bucy filter," *Journal of Statistical Physics*, vol. 118, pp. 145–176, January, 2005.

[22] S. Sinha and U. Vaidya, "Formalism for information transfer in dyamical networks," *IEEE Conference on Decision and Control*, 2015.

[23] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

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

[25] H. Marko, "The bidirectional communication theory- a generalization. of information theory," *IEEE Transaction on Communications*, vol. Com - 21, no. 12, pp. 1345–1351, Dec, 1973.

[26] T. Schreiber, "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461–464, July, 2000.

[27] L. Barnett, "Granger causality and transfer entropy are equivalent for gaussian variables," *Physical Review Letters*, vol. 103, pp. 238 701–1, Dec, 2009.

[28] I. J. Pérez-Arriaga, G. C. Verghese, and F. C. Schweppe, "Selective modal analysis with applications to electric power systems, part i: Heuristic introduction," *Power Apparatus and Systems, IEEE Transactions on*, no. 9, pp. 3117–3125, 1982.

[29] C. T. Lin, "Structural controllability," *Automatic Control, IEEE Transactions on*, vol. 19, no. 3, pp. 201–208, 1974.

[30] D. D. Siljak, *Decentralized control of complex systems*.   Courier Corporation, 2011.

[31] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints. ii. stabilization with limited information feedback," *Automatic Control, IEEE Transactions on*, vol. 44, no. 5, pp. 1049–1053, 1999.

[32] S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1549–1561, 2004.

[33] D. Li and N. Hovakimyan, "Bode-like integral for continuous-time closed-loop systems in the presence of limited information," *Automatic Control, IEEE Transactions on*, vol. 58, no. 6, pp. 1457–1469, 2013.

[34] S. Tatikonda and S. Mitter, "Control over noisy channels," *Automatic Control, IEEE Transactions on*, vol. 49, no. 7, pp. 1196–1201, 2004.

[35] N. Elia, "When bode meets shannon: Control-oriented feedback communication schemes," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1477–1488, 2004.

[36] J. A. Costanzo and J. Dunstan, "A survey of causality and directed information," 2014.

[37] Y.-H. Kim, H. H. Permuter, and T. Weissman, "Directed information and causal estimation in continuous time," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 819–823.

[38] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

[39] S. Sinha, U. Vaidya, and R. Rajaram, "Operator theoretic framework for optimal placement of sensors and actuators for control of nonequilibrium dynamics," *Journal of Mathematical Analysis and Applications*, vol. 440, no. 2, pp. 750–772, 2016.

[40] J. Ottino, *The Kinematics of Mixing: Stretching, Chaos, and Transport*. Cambridge, England: Cambridge University Press, 1989.

[41] G. Mathew, I. Mezić, S. Grivolopous, U. Vaidya, and L. Petzold, "Optimal control of mixing in stokes fluid flows," *Journal of Fluid Mechanics*, vol. 508, pp. 261–281, 2007.

[42] C. Kubrusly and H. Malebranche, "Sensors and controllers location in distributed systems - a survey," *Automatica*, vol. 21, no. 2, pp. 117–128, 1985.

[43] A. El Jai and A. J. Pritchard, *Sensors and controls in the analysis of distributed systems*. Halsted Press, 1988.

[44] W. Gawronski and K. Lim, "Balanced actuator and sensor placement for flexible structures," *International Journal of Control*, vol. 65, no. 1, pp. 131–145, 1996.

[45] U. Vaidya, R. Rajaram, and S. Dasgupta, "Actuator and sensor placement in linear advection pde with building system application," *Journal of Mathematical Analaysis and Applications*, vol. 394, pp. 213–224, 2012.

[46] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*. Cambridge, UK: Cambridge University Press, 1995.

[47] A. M. Mancho, D. Small, and S. Wiggins, "A tutorial on dynamical systems concepts applied to Lagrangian transport in oceanic flows defined as finite time data set: Theoretical and computational issues," *Physics Reports*, vol. 437, pp. 55–124, 2006.

[48] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor *et al.*, "Definition and classification of power system stability ieee/cigre joint task force on stability terms and definitions," *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1387–1401, 2004.

[49] T. Van Cutsem, "Voltage instability: phenomena, countermeasures, and analysis methods," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 208–227, 2000.

[50] W. Xu and Y. Mansour, "Voltage stability analysis using generic dynamic load models," *IEEE Transactions on Power Systems*, vol. 9, no. 1, pp. 479–493, 1994.

[51] Y. Wang, C. Wang, F. Lin, W. Li, L. Y. Wang, and J. Zhao, "Incorporating generator equivalent model into voltage stability analysis," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4857–4866, 2013.

[52] J. D. De Leon and C. W. Taylor, "Understanding and solving short-term voltage stability problems," in *Power Engineering Society Summer Meeting, 2002 IEEE*, vol. 2. IEEE, 2002, pp. 745–752.

[53] M. Paramasivam, A. Salloum, V. Ajjarapu, V. Vittal, N. B. Bhatt, and S. Liu, "Dynamic optimization based reactive power planning to mitigate slow voltage recovery and short term voltage instability," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3865–3873, 2013.

[54] E. G. Potamianakis and C. D. Vournas, "Short-term voltage instability: effects on synchronous and induction machines," *IEEE Transactions on Power Systems*, vol. 21, no. 2, p. 791, 2006.

[55] S. Dasgupta, M. Paramasivam, U. Vaidya, and V. Ajjarapu, "Real-time monitoring of short-term voltage stability using pmu data," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3702–3711, 2013.

[56] B. Gao, G. Morison, and P. Kundur, "Voltage stability evaluation using modal analysis," *IEEE Transactions on Power Systems*, vol. 7, no. 4, pp. 1529–1542, 1992.

[57] G. Verghese, I. Perez-Arriaga, and F. Schweppe, "Selective modal analysis with applications to electric power systems, part ii: The dynamic stability problem," *IEEE Transactions on Power Apparatus and Systems*, no. 9, pp. 3126–3134, 1982.

[58] P. W. Sauer and M. Pai, "Power system dynamics and stability," *Urbana*, vol. 51, p. 61801, 1997.

[59] L. Pereira, D. Kosterev, P. Mackin, D. Davies, J. Undrill, and W. Zhu, "An interim dynamic induction motor model for stability studies in the wscc," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 1108–1115, 2002.

[60] S. Sinha and U. Vaidya, "Causality preserving information transfer measure for control dynamical system," *IEEE Conference on Decision and Control*, pp. 7329–7334, 2016.

[61] ——, "On information transfer in discrete time dynamical system," *Indian Control Conference*, pp. 303–308, 2017.

[62] I. Dobson, H.-D. Chiang, J. S. Thorp, and L. Fekih-Ahmed, "A model of voltage collapse in electric power systems," in *Decision and Control, 1988., Proceedings of the 27th IEEE Conference on*. IEEE, 1988, pp. 2104–2109.

[63] V. Ajjarapu and B. Lee, "Bifurcation theory and its application to nonlinear dynamical phenomena in an electrical power system," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 424–431, 1992.

[64] I. J. Perez-Arriaga, G. C. Verghese, and F. C. Schweppe, "Selective modal analysis with applications to electric power systems, part 1: Heuristic introduction," *IEEE Transactions on Power Apparatus and Systems*, vol. 101, no. 9, pp. 3117–3125, 1982.

[65] S. Sinha, P. Sharma, U. Vaidya, and V. Ajjarapu, "Identifying causal interaction in power system : Information-based approach," *to appear in IEEE Conference on Decision and Control*, 2017.

[66] W. A. Hashlamoun, M. A. Hassouneh, and E. H. Abed, "New results on modal participation factors: Revealing a previously unknown dichotomy," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1439–1449, 2009.

[67] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Physical review letters*, vol. 75, no. 6, p. 1226, 1995.

[68] J. Buhl, D. J. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson, "From disorder to order in marching locusts," *Science*, vol. 312, no. 5778, pp. 1402–1406, 2006.

[69] R. Savit, R. Manuca, and R. Riolo, "Adaptive competition, market efficiency, and phase transitions," *Physical Review Letters*, vol. 82, no. 10, p. 2203, 1999.

[70] L. D. Landau, "On the theory of phase transitions," *Ukr. J. Phys.*, vol. 11, pp. 19–32, 1937.

[71] M. Baker and S. Glashow, "Spontaneous breakdown of elementary particle symmetries," *Physical review*, vol. 128, no. 5, p. 2462, 1962.

[72] E. Castellani, "On the meaning of symmetry breaking," *Symmetries in physics: Philosophical reflections*, pp. 321–334, 2003.

160

[73] R. T. Wicks, S. Chapman, and R. Dendy, "Mutual information as a tool for identifying phase transitions in dynamical complex systems with limited data," *Physical Review E*, vol. 75, no. 5, p. 051125, 2007.

[74] L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier, "Information flow in a kinetic ising model peaks in the disordered phase," *Physical review letters*, vol. 111, no. 17, p. 177203, 2013.

[75] S. N. Dorogovtsev, A. V. Goltsev, and J. F. Mendes, "Critical phenomena in complex networks," *Reviews of Modern Physics*, vol. 80, no. 4, p. 1275, 2008.

[76] G. Nicolis, I. Prigogine, and G. Nocolis, "Exploring complexity," 1989.

[77] S. Lang, "Algebra," *Graduate Texts in Mathematics*, vol. 1, no. 211, 2002.

[78] Y. Ueda, "Randomly transitional phenomena in the system governed by duffing's equation," *Journal of Statistical Physics*, vol. 20, no. 2, pp. 181–196, 1979.

[79] I. Kovacic and M. J. Brennan, *The Duffing equation: nonlinear oscillators and their behaviour*. John Wiley & Sons, 2011.

[80] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the atmospheric sciences*, vol. 20, no. 2, pp. 130–141, 1963.

[81] V. S. Afraimovich, V. Bykov, and L. P. Shilnikov, "On the origin and structure of the lorenz attractor," in *Akademiia Nauk SSSR Doklady*, vol. 234, 1977, pp. 336–339.

[82] J. Guckenheimer and R. F. Williams, "Structural stability of lorenz attractors," *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, vol. 50, no. 1, pp. 59–72, 1979.

[83] D. Ruelle, *Chaotic evolution and strange attractors*. Cambridge University Press, 1989, vol. 1.

[84] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *ACM SIGCOMM computer communication review*, vol. 29, no. 4. ACM, 1999, pp. 251–262.

[85] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, "The large-scale organization of metabolic networks," *Nature*, vol. 407, no. 6804, p. 651, 2000.

[86] P. Erdös and A. Rényi, "On random graphs, i," *Publicationes Mathematicae (Debrecen)*, vol. 6, pp. 290–297, 1959.

[87] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*. Cambridge university press, 1994, vol. 8.

[88] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *nature*, vol. 406, no. 6794, p. 378, 2000.

[89] B. Huang and U. Vaidya, "Data-driven approximation of transfer operators: Naturally structured dynamic mode decomposition," in *https://arxiv.org/abs/1709.06203*, 2016.

[90] C. Caramanis, S. Mannor, and H. Xu, "Robust optimization in machine learning," *Optimization for machine learning*, p. 369, 2012.

[91] M. Jovanovic, P. Schmid, and J. Nichols, "Low-rank and sparse dynamic mode decomposition," *Center for Turbulence Research Annual Research Briefs*, vol. 2012, pp. 139–152, 2012.

[92] N. Crnjaric-Zic, S. Macesic, and I. Mezic, "Koopman operator spectrum for random dynamical system," *arXiv preprint arXiv:1711.03146*, 2017.

[93] O. Junge, J. E. Marsden, and I. Mezic, "Uncertainty in the dynamics of conservative maps," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 2. IEEE, 2004, pp. 2225–2230.

[94] N. Takeishi, Y. Kawahara, and T. Yairi, "Subspace dynamic mode decomposition for stochastic koopman analysis," *Phys. Rev. E*, vol. 96, p. 033310, Sep 2017. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.96.033310

[95] S. Kutluay, A. Bahadir, and A. Ozde, "Numerical solution of one-dimensional burgers equation: explicit and exact-explicit finite difference methods," *Journal of Computational and Applied Mathematics*, vol. 103, no. 2, pp. 251 – 261, 1999.

[96] S. Sinha and U. Vaidya, "Causality preserving information transfer measure for control dynamical system," *IEEE COnference on Decision and Control*, pp. 7329–7334, 2016.

[97] ——, "On information transfer in discrete dynamical systems," *Indian Control Conference*, pp. 303–308, 2017.

[98] F. S. Mishkin, *The economics of money, banking, and financial markets*. Pearson education, 2007.

[99] G. M. Caporale, P. G. Howells, A. M. Soliman *et al.*, "Stock market development and economic growth: the causal linkage," *Journal of economic development*, vol. 29, no. 1, pp. 33–50, 2004.

[100] C. W. Granger, "Some recent development in a concept of causality," *Journal of econometrics*, vol. 39, no. 1-2, pp. 199–211, 1988.

[101] T. Vỳrost, Š. Lyócsa, and E. Baumöhl, "Granger causality stock market networks: Temporal proximity and preferential attachment," *Physica A: Statistical Mechanics and its Applications*, vol. 427, pp. 262–276, 2015.

[102] C. Hiemstra and J. D. Jones, "Testing for linear and nonlinear granger causality in the stock price-volume relation," *The Journal of Finance*, vol. 49, no. 5, pp. 1639–1664, 1994.

[103] T. Lux and M. Marchesi, "Volatility clustering in financial markets: a microsimulation of interacting agents," *International journal of theoretical and applied finance*, vol. 3, no. 04, pp. 675–702, 2000.

[104] S. Amari and H. Nagaoka, *Methods of information geometry*. American Mathematical Soc., 2007, vol. 191.

[105] D. B. Grusky, B. Western, and C. Wimer, *The great recession*. Russell Sage Foundation, 2011.

[106] R. E. Farmer, "The stock market crash of 2008 caused the great recession: Theory and evidence," *Journal of Economic Dynamics and Control*, vol. 36, no. 5, pp. 693–707, 2012.

[107] L. J. Christiano, M. S. Eichenbaum, and M. Trabandt, "Understanding the great recession," *American Economic Journal: Macroeconomics*, vol. 7, no. 1, pp. 110–67, 2015.

[108] A. Twin, "Dow sheds 486 points : Post-election worries about the weak economy are front and center." http://money.cnn.com/2008/11/05/markets/markets_newyork/, November 2008.