Comparing Dynamical Systems By Graph Matching Method

Jiongxuan Zheng

Math Department, Clarkson University Potsdam, NY-13676

Joseph D. Skufca

Math Department, Clarkson University Potsdam, NY-13676

Erik M. Bollt

Math Department, Clarkson University Potsdam, NY-13676

Abstract

In this paper, we consider comparing dynamical systems by using graph matching method either between the graphs representing the underlying symbolic dynamics, or graphs approximating the action of the systems on a fine but otherwise non generating partition. For conjugate systems, the graphs are isomorphic and we show that the permutation matrices that relate the adjacency matrices coincides with the solution of Monge's mass transport problem. We use the underlying Earth Mover's Distance (EMD) to generate the "approximate" matching matrix to illustrate the association of graphs which are derived from equal-distance partitioning of the phase spaces of systems. In addition, for one system which embeds into the other, we show that the comparison of these two systems by our method is an issue of subgraph matching.

Keywords: symbolic dynamics, graph matching, conjugate systems, isomorphic, Earth Mover's Distance, Monge-Kantorovich problem, embedding systems, subgraph matching.

1. Introduction

A basic question in science and in dynamical systems models of scientific problems is how to compare two systems. For an equivalence relation on dynamical systems, standard practice is to use the notion of *conjugacy*. The problem in practice is that it can be difficult in the general scenario of more than one dimension to find such a homeomorphic change of variables that relates the two systems, even if one exists. Furthermore if no such exact equivalence exists, the question becomes "is one system a good approximation of another?" In [3, 12], we explore this isuue by generating a change of variables between the systems, and then we measured the quality of the comparison by discussing the deviation of the resulting change of variables from a homeomorphism, which we called homeomorphic defect. The problem is the contraction mapping methods we used to identify the change of variables do not easily extend to the general scenario of multi-variate dynamical systems. Thus the need to develop alternative methods in the spirit of our goal [3, 12] of relaxing conjugacy to compare somewhat related dynamical systems. Here we will resort to a variant of the classic Monge-Kantorovich optimization problem [7, 8] to both built a useful change of variables and measure quality of the comparison through the underlying cost called the Wasserstein distance, in a spirit such as the related work for time series in [9].

In symbolic dynamics [4, 2, 1], a shift space is a set of infinite symbol streams representing the trajectory of a dynamical system. Shifts spaces with a finite set of forbidden blocks are called shifts of finite type. Such shifts spaces have a simple representation as a special case of so-called sophic shifts, that is a directed graph, where nodes are symbols. Furthermore we may define edges weights as the probability of a transition

Email addresses: `jizheng@clarkson.edu` (Jiongxuan Zheng), `jskufca@clarkson.edu` (Joseph D. Skufca), `bolltem@clarkson.edu` (Erik M. Bollt)

from one symbolic word to another. As such, the directed graph becomes a discrete representation of the Frobenius-Perron operator, [14] through an approximation method related to the Ulam's method, [15].

Within the usual dynamical systems framework, the determination of whether two systems are dynamically equivalent is based upon whether or not there is conjugacy between them, [2]. In terms of a symbolic dynamics representation [1, 4] of the underlying systems, the conjugacy matches symbol sequence from each symbolic space. System comparison within the concept of "conjugacy" would be straightforward, since all we need is to check is the existence of a similarity transformation that relates the adjacency matrices (one for each system) which generates the grammars of the underlying subshifts. Here we will demonstrate that such a matrix is equivalent to finding the Monge solution [8] to a specific mass transport problem. Because of well understood difficulties in finding Monge solutions, we instead choose to solve for the relaxed "matching matrix", which is actually a Kantorovich solution to relating the adjacency matrices of the dynamical systems. We use the earth mover's distance (EMD) [6] to generate such matching matrices which we show serves as a suitable commuter for the dynamical system. In addition, we prove that if one system embeds into the other, the comparison becomes a subgraph matching problem.

The outline of the paper is as follows: In section 2, we introduce the idea of our graph matching method to compare dynamical systems; In section 3, we give some theoretical results regarding the comparison between conjugate systems, with a particular example; In section 4, we consider the error analysis, with a numerical result to study the regularity property of our algorithm; in section 5, we extend our methods to compare systems where one embeds into the other, with both theoretical and numerical results given; In section 6, we show that our method can be easily extended to compare higher dimensional dynamical systems. We summarize our results in section 7.

2. Graph Representations of Dynamical systems

We adopt our notations from Lind and Marcus [1]: A graph G consists of a finite set $\nu = \nu(G)$ of vertices, and with a finite set $\varepsilon = \varepsilon(G)$ of edges. Each edge $e \in \varepsilon(G)$ starts at a vertex denoted by $i(e) \in \varepsilon(G)$ and ends at a vertex $t(e) \in \varepsilon(G)$. The weighted adjacency/stochastic matrix of G, denoted as A_G, is a conditional probability matrix where the (i, j) entry assigns the probability $P(j|i)$ (or $P(i \to j|i)$), the conditional probability from node i to j given the current state is in node i. The stochastic matrix is essentially a matrix description of a Markov chain on the graph $G = (\nu, \varepsilon)$ [1]. We note that some references may distinguish a stochastic matrix and the graph adjacency matrix, defining the latter as a binary matrix which assigns 1 to connected nodes, otherwise 0. In this paper, the transition matrix and stochastic matrix are used interchangeably, and (unless specifically stated otherwise) when we refer to "the adjacency matrix," we mean a *weighted* adjancency matrix with weights assigned to describe the appropriate transition probabilities.

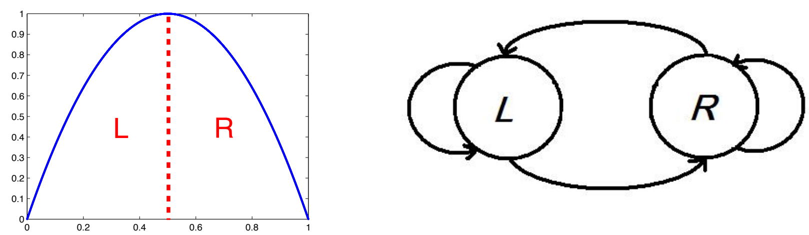
In practice, given a test orbit $\{x_k\}_{k=1}^N$ s.t. $x_{k+1} = f(x_k)$ from dynamical system f, then the transition probabilities can be estimated as follows,

$$P(j|i) \approx \frac{\#(\{x_k | x_k \in B_i \text{ and } f(x_k) \in B_j\})}{\#(\{x_k \in B_i\})} \tag{1}$$

as Ulam-Galerkin approximation of the Frobenius-Perron operator with a partition $\Gamma = \{B_i\}$, which is a finite family of connected sets with nonempty and disjoint interiors that covers the phase space [15, 17].

As example, Figure 1 shows a partition on the logistic map. We label the interval $[0, 1/2]$ with L, and $(1/2, 1]$ with R. Start from any initial point that yields a chaotic trajectory, record L or R in each iteration based on the evolution of dynamics. Then a weighted directed graph is built based on the recorded link list. In this example, the transition matrix would be a 2 by 2 matrix where entries are all 1/2.

The weighted directed graph is a simple representation of the dynamical system in that the adjacency matrix of the graph records all the allowable bi-infinite walks. Consequently, we assert that one way to relate the dynamics of two different systems is to perform a graph matching operation on the associated graph representations of the symbol dynamics. A key benefit of using a graph matching approach is that the study of graph matching techniques and algorithms is well developed in the field. Among many of the efficient graph matching methods, we use the Earth Mover's Distance, which essentially solves the discrete case of the Kantorovich's problem. What we hope to answer in this paper is,"what does the matching of the graphs using EMD tell us about the similarity of the dynamical systems?" As preliminary explanation, we briefly recall the Monge-Kantorovich problem.



(a) Logistic Map

(b) Graph for Logistic Map

Figure 1: Graph

2.1. Monge-Kantorovich Problem

The Monge-Kantorovich problem is an optimal transportation problem [8]: Specially, given two distributions μ for space X and ν for Y, one is required to find the optimal transportation plan within a given some metric describing the transportation cost. The Monge's problem is formalized as:

$$Minimize\ \ I[T] := \int_X c(x,\ T(x)))d\mu(x)$$

for all measurable map $T : X \to Y$ such that ν is a push-forward [8] of μ by T. The term $c(x,\ T(x))$ defines the cost to transport a unit from x to $T(x)$.

On the other hand, the Kontorovich's problem is a relaxed version of the Monge's one. It allows each point in X to be associated to different points in Y. We can no longer use a map T to describe this association. We shall use a probability measures π, the transference plan, on the product space $X \times Y$, where informally $d\pi(x, y)$ measures the amount of mass transferred from location x to y. We denote the set of all such probability measures by

$$\Pi(\mu, \nu) = \{\pi \in P(X \times Y) : \pi[A \times Y] = \mu[A],\ \pi[X \times B] = \nu[B]\ for\ all\ measurable\ A, B\}$$

And the Kontorovich's optimal transportation problem can be formalized as follow:

$$Minimize\ \ I[\pi] := \int_{X \times Y} c(x, y)d\pi(x, y)$$

for $\pi \in \Pi(\mu, \nu)$, where $\Pi(\mu, \nu)$ (defined above) is called the space of transference plans from X to Y.

In the later discussion of this paper, we focus on the discrete version of the Monge-Kantorovich problem, with solution given as a matrix F that expresses the optimal transportation in that matrix entry f_{ij} gives the amount of mass to be moved from discrete location x_i to location y_j. In that setting, the special case of a Monge solution would then be a permutation matrix, because the Monge problem requires that the full mass at location x_i be moved to a single site y_j.

We will compare two systems by studying the Monge-Kantorovich solution between the first eigenvectors of their associated adjacency matrices relying on Ulam's conjecture. Ulam's conjecture was proposed by S. Ulam [15] and referred that Ulam-Gelerkin matrix determined by Eq. 1 has dominant state/eigenvector that converges weakly to invariant measure/distribution of the Frobenius-Perron operator, which has been proven for certain special dynamical systems.

3. Analysis

In this section, we relate graph matching and conjugacy. Particularly, we prove an interesting relationship between the matching matrix F that describes the Monge solution to a particular problem, and the adjacency matrices, A_{G_1} and A_{G_2}, that describe the dynamics of the two systems:

$$F \cdot A_{G_1} = A_{G_2} \cdot F.$$

We first introduce the notation: A partition $\Gamma = \{x_0,\ x_1, \cdots,\ x_m\}$ on interval $[a, b]$ is a set of points where $a = x_1 < x_1 < \cdots < x_m = b$. We denote $|\Gamma|$ as the number of subintervals after partitioning, which we call the cardinality of partition. Let G be a graph with edge set ε and adjacency matrix A_G. The edge shift X_G or X_A is the shift space over the alphabet $\mathscr{A} = \varepsilon$ specified by $X_G = X_A = \{\xi = (\xi_i)_{i\in\mathbb{Z}} \in \varepsilon^{\mathbb{Z}} : t(\xi_i) = i(\xi_{i+1})$ for all $i \in \mathbb{Z}\}$. The shift map on X_G or X_{A_G} is called the edge shift map and is denoted by σ_G or σ_{A_G}.

3.1. Conjugate Systems

Two dynamical systems are said to be *topologically conjugate* if there exists a homeomorphism which describes the change of coordinates between systems. Consequently, when we associate these systems to dynamics on graphs, we may use that homeomorphism to relate the partition in one system to the partion in the other. Because trajectories in each system are matched (via the homeomorphism), allowable paths through the graphs must also be matched, and the graphs of the two systems would be. So in this case, the associated graph adjacency matrices, say A (for one system) and B (for the other), are *similar,* and related by a permutation matrix P. Essentially, P tells how the nodes of the compared graphs are associated, in a 1-to-1 way, so that $A = P^{-1}BP$. In addition, P relates the eigenvectors of A and B as a linear isomorphism [1]. And, we can identify conjugate edge shifts from these conjugate systems (Theorem 4 in Appendix), which also determines a so-called "strong shift equivalent" relationship. More precisely, we have the following definition:

Definition 1 ([1]). *Let A and B be nonnegative integral matrices. An elementary equivalence from A to B is a pair (R, S) of rectangular nonnegative integral matrices satisfying*

$$A = RS \text{ and } B = SR$$

In this case we write $(R, S) : A \cong B$. A strong shift equivalence of lag l from A to B is a sequence of l elementary equivalences

$$(R_1, S_1) : A = A_0 \cong A_1,\ (R_2, S_2) : A_1 \cong A_2 \cdots (R_l, S_l) : A_{l-1} \cong A_l = B$$

In this case we write $A \approx B$ (lag l). We say that A is strong shift equivalent to B (and write $A \approx B$) if there is a strong shift equivalence of some lag from A to B.

In the following we show that, for isomorphic graphs, which we obtain from conjugate systems by an appropriate partitioning, the permutation matrix P which provides the similarity transform between the associated adjacency matrices is the same matrix identified by the Monge solution to an appropriately defined transport problem.

Theorem 1. *Suppose we are given two conjugate systems SYS_1 and SYS_2, and partitions Γ_i, $i = 1, 2$, with $|\Gamma_i| = M$, such that the resulting edge shifts X_{G_1} is conjugate to SYS_1, X_{G_2} is conjugate to SYS_2. Let A_{G_1} be the adjacency matrix of the graph generated from SYS_1 with this partition, and similarly A_{G_2} for SYS_2. Suppose A_{G_1} is isomorphic to A_{G_2}. Let F be the Monge solution between dominant eigenvectors of adjacency matrices [2] of graphs from SYS_1 and SYS_2, then F satisfies*

$$F \cdot A_{G_1} = A_{G_2} \cdot F.$$

[2]

Remark 1. Note that in the proof, we apply theorems and definitions from the appendix. Some of the fundamental theorems and definitions that we are using require matrices to be over $\mathbb{Z}$ or $\mathbb{Z}^+$. But as we define before in section 2, we generally consider adjacency matrices to be stochastic matrices. The stochastic matrix setting makes our method easy to be implemented, as will be illustrated in section 3.3. Because these stochastic weights are determined by counting transitions in a long trajectory, and then normalizing each row (to row sum one), each of these stochastic matrices has rational entries, and appropriate scaling yields an integer matrix. This renormalized matrix is similarly interpretable as a weighted adjacency matrix. Consequently, the proofs that rely on integer matrix theorems remain valid under such a renormalization. The Monge's solution would be interpreted as showing how to move "counts" instead of "densities." Even though some references [1, 16] discuss shift equivalence issues for irreducible shift of finite type with a Markov measure, we prefer not to introduce additional theorems and definitions to address this technical detail, relying simply on the frequency counting to justify our approach.

Proof. Since SYS_1 is conjugate to SYS_2, $X_{G_1} \cong X_{G_2}$. By Theorem 4, $A_{G_1} \approx A_{G_2}$ (strong shift equivalent). By Definition 5 and Theorem 5, $A_{G_1} \sim A_{G_2}$ (shift equivalent). By Definition 6 and Theorem 6, $A_{G_1} \sim_{\mathbb{Z}} A_{G_2}$ (shift equivalent over $\mathbb{Z}$). On the other hand, since A_{G_1} is isomorphic to A_{G_2}, there exists a permutation matrix P such that $A_{G_1} = P^{-1}A_{G_2}P$. Now if we let $S = P$, $R = P^{-1}A_{G_2}$, then $A_{G_1} \sim_{\mathbb{Z}} A_{G_2}(lag\ 1)$ according to Definition 1. By Theorem 7 and Theorem 8, $w = Sv \neq 0$ where v and w are Perron eigenvectors of A_{G_1} and A_{G_2}, respectively, and S satisfies $S \cdot A_{G_1} = A_{G_2} \cdot S$. In addition, S is a permutation matrix, hence S is invertible.

On the other hand, let F be the Monge solution between v and w, where v and w are dominating eigenvectors of A_{G_1} and A_{G_2}, then F satisfies $w = Fv \neq 0$. We claim that $F = S$.

By *Choquet's theorem* and *Birkhoff's theorem* [8], F is a permutation matrix, which implies F is invertible. Substitute $v = F^{-1}w$ into $w = Sv$, we have

$$w = S \cdot F^{-1}w, \tag{2}$$

where $S \cdot F^{-1}$ is also always a permutation matrix, regardless of the cardinality of partition. We note that (2) is not sufficient in that $S \cdot F^{-1}$ could be a rotation matrix other than identity matrix that satisfies $w = S \cdot F^{-1}w$.

We prove by induction that, $S \cdot F^{-1}$ is an identity matrix. In the following, we let $E_N = S \cdot F^{-1}$, where the cardinality of partition $M = N$, where we intent to use induction on N.

When $N = M = 2$, the only permutation matrices that $S \cdot F^{-1}$ can take are

$$E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

or

$$\hat{E}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

but $|\hat{E}_2| = -1$, so $\hat{E}_2$ can not be a rotation matrix. Thus $S \cdot F^{-1} = E_2$.

Now we suppose when $M = N$, E_N is an N by N identity matrix, so that $w = S \cdot F^{-1}w$.

When $M = N + 1$, in order that the resulting edge shifts with cardinality of partition $N + 1$ is conjugate to the one with cardinality of partition N, the graph with $N + 1$ nodes has to be a splitting of the one with N nodes [1]. Without loss of generality, we can assume the first node of the graph splits into two nodes, from cardinality of partition N to $N + 1$. Then the first bin of dominate eigenvector w of case N would break into two for case $N + 1$, with the rest of bins unchanged. For example, from Figure 2(a) to Figure 2(b), *bin* 1 breaks into *bin* 1*a* and *bin* 1*b*.

Regardless of how *bin* 1*a* and *bin* 1*b* match each other, we claim that the rest bins of w, like bins from 2 to 5 in Figure 2(b), match themselves identically as before. Because otherwise, if we amalgamate *bin* 1*a* and *bin* 1*b* back into *bin* 1, or say from case $N + 1$ back to case N, the bins of w would not match themself identically. This contradicts to our assumption of the matching of bins of w when $M = N$.

Also, since $S \cdot F^{-1}$ is a permutation matrix, E_{N+1} would has the forms of

$$E_{N+1} = \begin{bmatrix} E_2 & 0 \\ 0 & E_{N-1} \end{bmatrix}$$

or

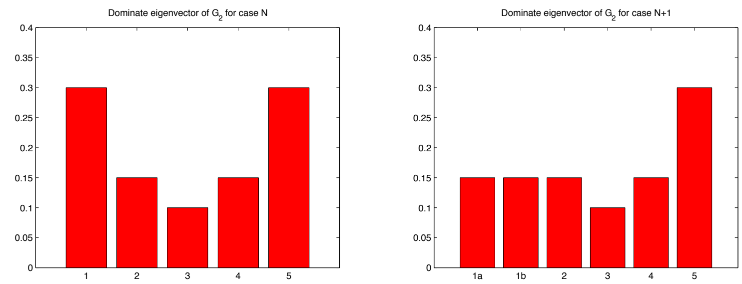
$$\hat{E_{N+1}} = \begin{bmatrix} \hat{E}_2 & 0 \\ 0 & E_{N-1} \end{bmatrix}$$

where E_{N-1} is a $N-1$ by $N-1$ identity matrix. Thus, by the calculation of determinant of block matrix, we know that $|\hat{E_{N+1}}| = -1$, which can not be a rotation matrix. Thus, $S \cdot F^{-1} = E_{N+1}$, which is a $N + 1$ by $N + 1$ identity matrix.

Similar for the case $v = F \cdot S^{-1}v$. Thus, we have $S = F$.

So we can consider finding the similarity matrix P as a process of choosing the association matrices that minimize the transportation cost.

However, there is not a general algorithm to find the similarity matrix P. Although we prove the equivalence between finding the similarity matrix and finding Monge's solution, that equivalence does not mean that either problem is easily solved. In fact, Monge's problem is difficult because the metric that defines the



(a) Dominate eigenvector w for case N (b) Dominate eigenvector w for case $N + 1$

Figure 2: Dominate eigenvectors of graph for case N and $N + 1$

topology of the transport problem could be very degenerate from the point of view of convexity properties [8]. For non-conjugate systems (and non-isomorphic graphs), there still exists a Monge solution which associates nodes of the two graphs in a 1-to-1 fashion in a way that minimizes the "cost of association" with respect to a particular choice of norm to measure the transportation cost. We observe that the Monge solution yields a linear transformation from the dominate eigenvector of one graph to the other. On the other hand, the Monge solution, does *not* provide a similar linear transformation between the adjacency matrices when the underlying graphs are not isomorphic. Consequently, the relaxed version, i.e. Kantorovich's problem, provides a more natural and implementable method to find an "optimal matching", where optimality is with respect to the relaxed transport problem.

3.2. Partitioning Issues

As additional reasons for framing this "system matching" as a Kantorovich problem is that framing as a Monge problem requires that we our partition is such that the edge shift is conjugate to the original dynamics. In most practical cases, it is hard or costly to find an appropriate partition (i.e. Markov partition) so that the resulting graph can fully describe the original dynamical system. Rather, we seek a method that can be applied to comparing general dynamical systems *without* solving the "hard" problem of finding a Markov partition. As alternative, we will use an equal-distance partition. We have to be very careful here, because in general the equal-distance partition won't give a conjugate shift space. So, in principle, if we proceed our method by comparing dynamical systems via equal-distance partition, we are actually comparing different systems than what we really would like to compare. However, this issue can be addressed as an "approximation problem." If we refine our equal-distance partition to sufficiently small cells, then the resulting shift space can approximate the one that would be observed from a Markov partition.

As example of this approximation via grid refinement, L. Billing and E. Bollt proved that the family of skew tent map is Markov for a dense set of parameters in the chaotic region [10]. In particular, in any given neighborhood of a non-Markov map in this chaotic region, there exists a Markov map that uniformly approximates the non-Markov one. J. Zheng proved that this approximation implies approximation in the sense of conjugacy [12]. In addition, H. Teramoto and T. Komatsuzaki showed that if a Markov partition has a certain relationship called "map-refinement of the other Markov partition", the shift spaces corresponding to these two Markov partitions are topologically the same [11]. Thus, an adequately refined equidistant partition can approximate a Markov partition in the sense of conjugacy. In section 4.1, we examine this numerical error estimate for a tent map example.

3.3. Graph Matching Method for Comparing Dynamical Systems by Earth Mover's Distance

There are several cases where the Kantorovich's solution coincides with the Monge's solution. When we consider the structure of the space with metric, or say the cost function, defined with a regular $p > 1$ norm, the strict convexity of the metric guarantees there is a unique solution to the Kantorovich problem, which turns out to be also the solution to the Monge problem [8].

However, because of the partitioning issues that we have discussed before, the graphs representations of the systems are somehow an "approximation". So the adjacency matrices from conjugate systems are no longer similarly related due to inexact partition of phase space. However, we seek to relax the problem in the way of allowing partial matching, which is to find the Kantororich's solution, to describe the association.

In probability theory, the Earth Mover's Distance (EMD) is a measure of distance between two probability distributions. EMD can be viewed as the solution of the discrete case of Kantorovich problem. It is a bipartite network flow problem which can be formalized as a linear programming problem [6]: Let I be the set of supplies, J be the set of consumers, and c_{ij} be the cost to ship a unit from $i \in I$ to $j \in J$. We want to find a set of flow (matching matrix) f_{ij} to minimize the overall cost:

$$\sum_{i \in I} \sum_{j \in J} c_{ij} f_{ij},$$

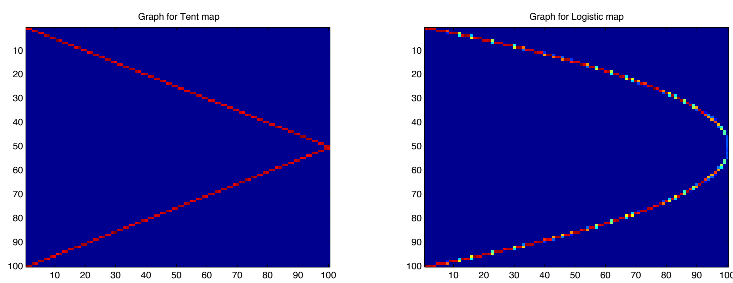
subject to the constrains:

$$\begin{gathered} f_{ij} \geq 0, \ i \in I, \ j \in J \\ \sum_{i \in I} f_{ij} = y_j, \ j \in J \\ \sum_{j \in J} f_{ij} \leq x_i, \ i \in I \end{gathered}$$

where x_i is the total supply of supplier i, and y_j is the total capacity of consumer j.

Here we consider the cost function $c_{ij} = |x - y|^2$, with the metric to be the regular 2 norm. In this case, the optimal matching/flow is the gradients of the convex functions, which are monotone and orientation preserving. On the other hand, we consider the matrix norm to be the induced 2 norm, since with induced norm, the matrices can be viewed as operators. And if two systems are conjugate, they are isomorphic as linear transformations if we restrict the operators/matrices to their eventual range [1].

We give the example about how to create the matching matrix between systems: suppose we are given two dynamical systems, i.e. the logistic map and tent map, we make a equidistant partition on the phase space with partitionning numbers M and N, respectively. We randomly choose an initial starting point, and then iterate it on each system, record the interval/node that the trajectory passes, and built the graphs based on the recording link list. Figure 3 presents the graphs of logistic map and tent map.



(a) Adjacency Matrix of Tent Map (b) Adjacency Matrix of Logistic Map

Figure 3: Graph

The matching matrix (Figure 4) generated by Earth Mover's distance associates the nodes between tent graph and logistic graph in the way of minimizing the overall cost determined by predefined topology. Notice the matching itself is a matrix. In Figure 4, the x-axis are nodes for tent graph, y-axis are for logistic graph. Color intensity implies similarity of the matching. We also plot the homeomorphsim function of tent and logistic in blue curve. We can see that both the matching matrix and the conjugacy describe the dirt of the tent distribution is moved in the direction so that it can match the logistic map distribution.

3.4. Further Discussion about the Measure on Adjacency Matrices

In Figure 4, we present the matching matrix between tent map and logistic map, which are conjugate systems. Note we apply the Earth Mover's Distance on their adjacency matrices whose entries are assigned the conditional probability measure. We can think of the matching matrix as a different representation of conjugacy, since the EMD is actually giving the transformation between the first eigenvectors of the tent map and logistic map.

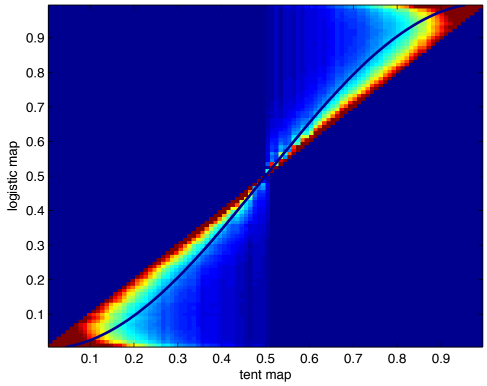


Figure 4: Graph matching between Logistic map and Tent map.

This matching method may be confusing sometimes, since there are systems that are not conjugate but generating a "conjugate like" matching matrix. For example, if we are comparing a two hump tent map, instead of the one hump tent map, to the logistic map, our method would still give a matching matrix as shown in Figure 4. This comes from the fact that, different systems may have the same invariant measures.

Thus we also present an alternative but similar approach for our matching method. Instead of assigning the conditional probability measure to the entries of the adjacency matrices of underlying systems, we equalize these non-zero weights of each row of the adjacency matrices.

Figure 5 shows the matching matrix between the tent map and the logistic map with equalized weights on their adjacency matrices. This focuses more on the topological relation between systems. In the following section, we will discuss regularity properties when using the adjacency matrices with probability measure.

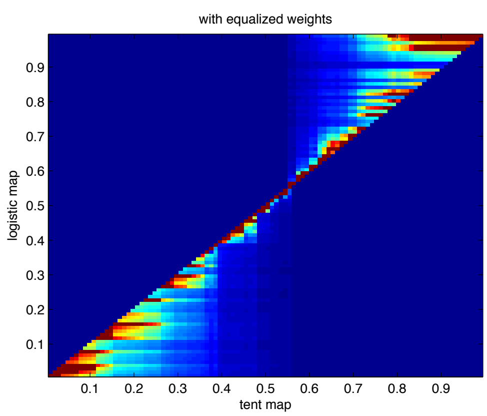


Figure 5: Matching Matrix using Equalized weights on the Adjacency Matrices

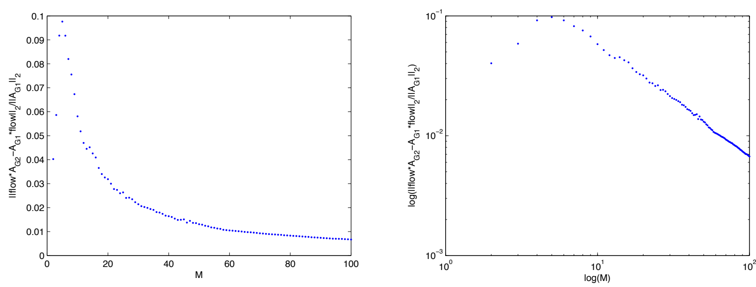
4. Error Analysis and Regularity Property Analysis

The graph matching method to compare dynamical systems that we discussed in the previous section suffers from both numerical and theoretical errors: the partition errors that we mentioned in section 3.2; the statistical errors from comparing the graphs which are constructed by a particular finite trajectory; the linear program solver error from the Earth Mover's distance algorithm, which can be reduced by resetting the stopping criteria stricter. In this section, we will show, numerically, that these errors can be reduced by refining the partition and increasing sampling size.

On the other hand, we will consider the regularity of our method, i.e. we will show whether "closer" systems gives a smaller number of "dissimilarity" for matching. As our goal of this work is to compare general dynamical systems by graph matching method. We seek to study the metric of "dissimilarity" as the measure of distance from being conjugacy, or isomorphism for the associate graphs.

4.1. Partition Error

As we have discussed the partitioning issue in section 3.2, we can use the equidistant partition to approximate the Markov partition with more and more cardinality of partition M. Continuing the example in previous section, we consider the logistic map T_1 and the symmetric full tent map T_2, and the value $\frac{\|flow*A_{T2}-A_{T1}*flow\|_2}{\|A_{T1}\|_2}$, a relative error of being isomorphic, when increasing M. Figure 6 suggests the error is decreasing approximately polynomially as M increases. In particular, when cardinality of partition is 100, i.e. $M = 100$, $\frac{\|flow*A_{T2}-A_{T1}*flow\|_2}{\|A_{T1}\|_2} = 0.46\%$.



(a) Partition error for refining partition

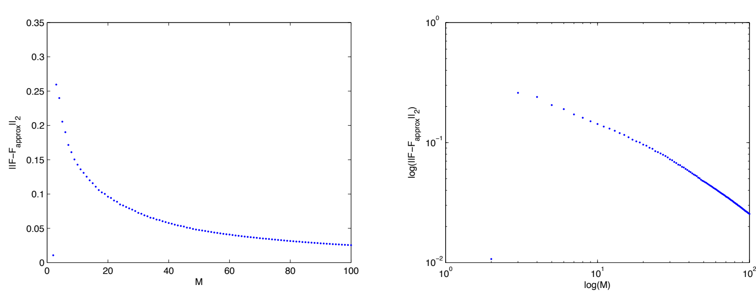
(b) Partition error for refining partition in *log* scale

Figure 6: Partition error $\frac{\|flow*A_{T2}-A_{T1}*flow\|_2}{\|A_{T1}\|_2}$ with number of partition M increasing.

4.2. Statistical Error

We apply the Earth Mover's distance to compare graphs which are from taking a particular finite trajectory of systems. If we let F be the matching matrix between the exact graphs from systems, which can be solved analytically in this example, and let $F_{approximate}$ be the matching matrix between the approximate graphs from finite trajectories, and $\|F - F_{approximate}\|_2$ be the error. Then by taking longer iterations for the trajectory and transient, we have a more precise statistical results, which gives a smaller number of error.

On the other hand, as shown in Figure 7, if we let the cardinality M increase, the error is decreasing approximately polynomially.



(a) Error analysis for refining partition

(b) Error analysis for refining partition in *log* scale

Figure 7: Error analysis for refining partition

4.3. Regularity Property for the Metric

We have discussed the issue for conjugate systems, which allows perfect matching for the corresponding graphs. What we are going to consider is, whether "closer" systems implies smaller measure of the dissimilarity, which we regard as a regularity issue of the method. Here we consider an example with a logistic map L and a symmetric full tent map with additive noise T. We let the measure of dissimilarity to be $\|F * A_T - A_L * F\|_2$, where F is the matching matrix by Earth Mover's distance. As shown in Figure 8, when we decrease the amplitude of the additive noise, the dissimilarity value $\|F * A_T - A_L * F\|_2$ is decreasing to a certain value monotonously.

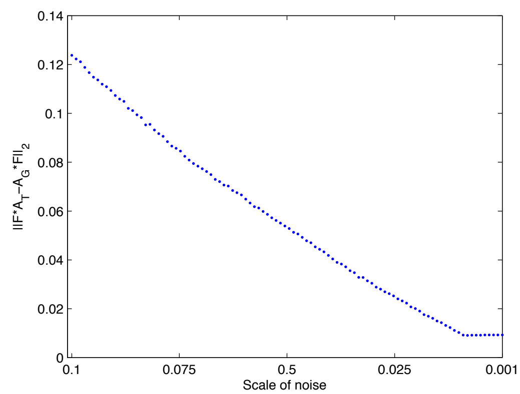


Figure 8: Measure of dissimilarity $\|flow * A_T - A_L * flow\|_2$ with decreasing scale of noise for the noisy tent map.

5. For Embedding Systems

We have established the graph matching method to compare conjugate systems. However, in most of the cases, the underlying systems are not perfectly homeomorphic. It is our goal of this topic to develop the graph matching method to compare general dynamical systems. But the difficulty is, even we can follow the same procedure to present the matching matrix like Figure 4 for general non-conjugate systems, the theoretical explanation and justification stand in our way. But at this stage, we have shown that, if one system embeds into the other, the issue would become a subgraph matching problem.

We have the following definitions regarding to embedding:

Definition 2. *If $\phi : X \to Y$ is one-to-one, then ϕ is called an embedding of X into Y.*

Definition 3. *$P(X)$ denote the set of all periodic points in X. $q_n(X)$ is the number of points with least periodic n, i.e. $q_n(X) = |\{x : \phi^n(x) = x,\ \phi^k(x) \neq x,\ 0 < k < n\}|$.*

Note that if $\phi : X \to Y$ is an embedding, then ϕ restricts to a one-to-one shift-commuting mapping from $P(X)$ into $P(Y)$. The existence of such a mapping is equivalent to the condition $q_n(X) \leq q_n(Y)$ for all $n \geq 1$.We denote these equivalent conditions by $P(X) \hookrightarrow P(Y)$, and call this the embedding periodic point condition.

In addition, we refer "the induced subgraph" as:

Definition 4. *Let H be a graph with vertex set v. For each subset w of v define the induced subgraph of H from w to have vertex set w and edge set the collection of all edges in H that start and end in w. An induced subgraph of H is one that is induced by some subset of v,*

Note that if H has adjacency matrix B, and G is an induced subgraph of H, then A_G is a principal submatrix of B; i.e., A_G is obtained from B by deleting the jth row and j column for a certain set of j's.

We would first give the proof of Theorem 2, which is similar to the proof of *Masking Lemma* (Theorem 11). But the difference is, our proof is like "cutting edges", while the proof of *Masking Lemma* is "adding edges".

Theorem 2. *Let G and H be graphs. Suppose that X_G embeds into X_H. Then there is a induced subgraph K of H, such that $X_G \cong X_K$.*

Proof. Let G' and H' be graphs constructed in Theorem 10. Then $A_H \cong A_{H'}$. So there is a sequence of graphs

$$H' = H_0,\ H_1 \cdots,\ H_k = H$$

such that $A_{H_i} \cong A_{H_{i+1}}$ for $0 \leq i \leq k-1$. By the next Lemma that will be proved, we will find a sequence of graphs

$$G' = G_0,\ G_1 \cdots,\ G_k$$

such that each G_i is an induced subgraph of H_i, and the elementary equivalence from A_{H_i} to $A_{H_{i+1}}$ extends to one from A_{G_i} to $A_{G_{i+1}}$. We set $K = G_k$. Then $A_K \cong A_{G'} \cong A_G$ and K is an induced subgraph H.

Lemma 1 ("cutting edges"). *Let $G_1,\ H_1,\ H_2$ be graphs such that $A_{H_1} \cong A_{H_2}$ and G_2 is an induced subgraph of H_1. Then there is a graph G_2 such that $A_{G_1} \cong A_{G_2}$ and G_2 is an induced subgraph of H_2.*

Proof. Let $A = A_{H_1}$ and $B = A_{H_2}$. Let $(R, S) : A \cong B$ so that $RS = A$ and $SR = B$. Let $G_{R,S}$ be the auxiliary graph. This graph has vertex set which is the disjoint union of $v(H_1)$ and $v(H_2)$. It contains a copy of H_1, whose edges are called $A - edges$, and a copy of G_2, whose edges are $B - edges$. For each $I \in v(H_1)$ and $J \in v(H_2)$, there are R_{IJ} edges from I to J, called $R - edges$, and S_{JI} edges from J to I, called $S - edges$.

We use G_1 to reduce $G_{R,S}$, forming a new graph $\bar{G}$ as follows. Let $v = v(H_1) \setminus v(G_1)$ and $\varepsilon = \varepsilon(H_1) \setminus \varepsilon(G_1)$. Subtract $v \cup \varepsilon$ to the vertex set of $G_{R,S}$ to form the vertex set of $\bar{G}$. To avoid notational confusion, let I_e denote the subtracted vertex corresponding to $e \in \varepsilon$. It will be helpful to think of vertices in v as being subtracted to the "$H_1 - part$" of $G_{R,S}$. And vertices I_e, for each $e \in \varepsilon$, as being subtracted to the "$H_2 - part$" of $G_{R,S}$. For each $e \in \varepsilon$, subtract a $R - edges$ $r(e)$ from the initial state of e to I_e, and $S - edge$ $s(e)$ from I_e to the terminal state of e. For each $R, S - path$ from I to I' that this cuts, subtract to $\bar{G}$ an $A - edge$ from I to I'; similarly, for each $S, R - path$ from J to J' that this cuts, subtract to $\bar{G}$ a $B - edge$ from J to J' (see Figure 9). This completes the construction of $\bar{G}$.

Each $e \in \varepsilon$ determines a cut $R, S - path$ $r(e)s(e)$ from $G_{R,S}$ with the same initial and terminal states in G_1 as in $\bar{G}$, and these are the only subtracted $R, S - paths$. Hence the graph with vertices $v(H_1) \setminus v$ and $A - edges$ of $\bar{G}$ is isomorphic to G_1. Let G_2 be the graph with vertices $v(H_2) \setminus \{I_e : e \in \varepsilon\}$ and all $B - edges$ of $\bar{G}$. Let $\bar{R}$ and $\bar{S}$ record the incidences of $R - edges$ and $S - edges$ in $\bar{G}$. The $\bar{G}$ is the auxiliary graph $G_{\bar{R},\bar{S}}$ corresponding to the elementary equivalence $(\bar{R}, \bar{S}) : A_{G_1} \cong A_{G_2}$. Furthermore, G_2 is the subgraph of H_2 induced from the vertex subset v_{G_2}.

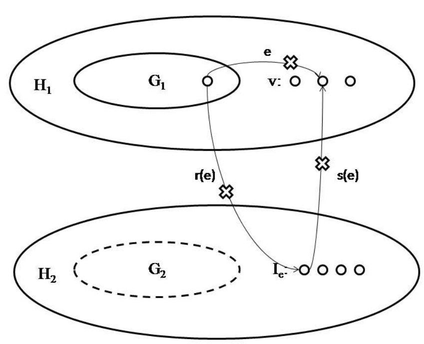


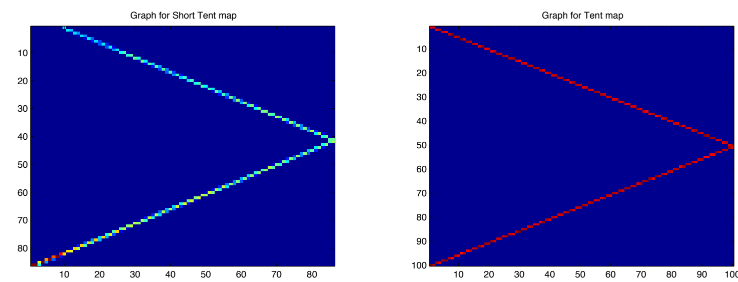
Figure 9: Auxiliary graph

Combine Theorem 2 and Theorem 1, we have the following theorem, which establish the theory for non-conjugate systems where one embeds into the other.

Theorem 3. *Suppose we are given two non-conjugate systems SYS_1 and SYS_2, and a partition Γ with $|\Gamma| = M$, such that the resulting edge shifts X_{G_1} is conjugate to SYS_1, X_{G_2} is conjugate to SYS_2, where A_{G_1} is the adjacency matrix of graph that generated from SYS_1 with this partition, and A_{G_2} is from SYS_2. Suppose X_{G_1} embeds into X_{G_2}. Then there is a induced subgraph G'_2 of G_2, such that there is a matrix F (the Earth Mover's Distance matching matrix between graphs from X_{G_1} and $X_{G'_2}$), satisfies*

$$F \cdot A_{G_1} = A_{G'_2} \cdot F.$$

A simple illustration would be the short symmetric tent map and full tent map, which are not conjugate systems, and the short one embeds into the full one.



(a) Adjacency Matrix for short tent map (b) Adjacency Matrix for full tent map

Figure 10: Induced Subgraph

Figure 10(a), denote A_G, is the graph for short symmetric tent map, whose peak is at 0.95, while Figure 10(b), denote A_K, is the graph for short symmetric tent map. We claim that Figure 10(a) is a induced subgraph of Figure 10(b), or say Figure 10(a) can be enlarged to Figure 10(b).

6. For Higher Dimensional Dynamical Systems

The graph matching method for higher dimensional dynamical systems can be easily extended from the 1-d case. For 2-d dynamical systems, we partition the phase by equal sized boxes. Then the graphs are determined by investigating the trajectory of a randomly chosen starting point, like the procedure for 1-d maps. And EMD can be applied to obtain the matching matrix.

As a example, we use the graph matching method to compare two conjugate Henon map (blue and red Henon maps in Figure 11). We choose the blue Henon map as:

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n \\ y_{n+1} = bx_n \end{cases} \tag{3}$$

with parameters $a = 1.4$, $b = 0.3$, and the red Henon map, which is a cubic transformation of the blue one, i.e. $x_{red} = x_{blue}^3$; $y_{red} = y_{blue}^3$, which is a smooth change of coordinate.

Figure 12(a) and Figure 12(b) show the graphs of the blue Henon and red Henon, respectively. Applying EMD on this two graphs gives us the matching matrix Figure 12(c). Note that due to the coarse partition, which is good for visualization, the matching matrix only suggests a rough association between the boxes in Figure 11.

7. Conclusion

In this paper, we compare dynamical systems by using graph matching method. For conjugate systems, which have isomorphic graphs, we show that the permutation matrix that relates the adjacency matrices of the graphs is the same as the solution of Monge's problem. We use Earth Mover's Distance (EMD) to generate the "approximated" matching matrix to illustrate the association of graphs which are derived from equal-distance partitioning of the phase spaces of systems. In addition, for one system embedding into the other,

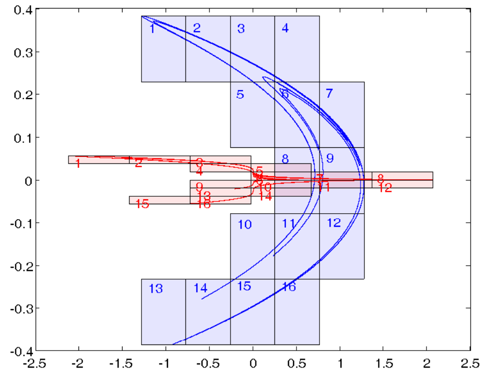
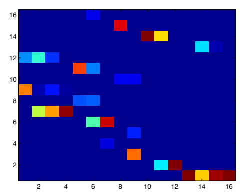
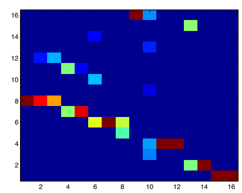


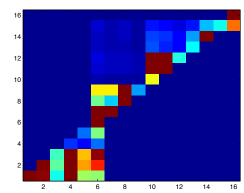
Figure 11: Conjugate Henon Maps



(a) Blue Henon Maps



(b) Red Henon Map



(c) Matching Matrix

Figure 12: Graph Matching for 2D Dynamical Systems

we show that the comparison of these two systems by our method is an issue of subgraph matching. In the end we show how to extend the method to compare higher dimensional systems.

For our ultimate objective and future work, we seek to theoretically extend the graph matching method to compare more general dynamical systems, which are not necessary conjugate or embedding. On the other hand, the regularity property of our method should be justified. In particular, "closer" systems in the metric of graphs' distance should imply closer dynamical relationships, as considered in paper [3]. And by this setting, finding an optimal system to model a time series data or approximate a system can be achieved by minimizing the "dissimilarity" of matching.

Acknowledgments

This work was accomplished with funding provided by National Science Foundation grant DMS-0708083

.

Appendix

Here we list the theorems and definitions that we used in our proof. The reader can refer to Lind [1] for further details.

Theorem 4 ([1]). *The edge shifts X_A and X_B are conjugate if and only if matrices A and B are strong shift equivalent, we write*

$$X_A \cong X_B \Longleftrightarrow A \approx B$$

Definition 5 ([1]). *Let A and B be nonnegative integral matrices and $l \geq 1$. A shift equivalence of lag l from A to B is a pair (R, S) of rectangular nonnegative integral matrices satisfying the four shift equivalence equations*

$$\begin{aligned}&(i) AR = RB, SA = BS\\&(ii) A^l = RS, B^l = SR\end{aligned} \tag{4}$$

Denote by $(R, S) : A \sim B$ (lag l). We say that A is shift equivalent to B, written $A \sim B$, if there us a shift equivalence from A to B of some lag.

Theorem 5 ([1]). *Strong shift equivalence implies shift equivalence, we write*

$$A \approx B \Longrightarrow A \sim B.$$

Definition 6 ([1]). *Let A and B be integral matrices. Then A and B are shift equivalent over $\mathbb{Z}$ with lag l, written $A \sim_{\mathbb{Z}} B$ (lag l), if there are rectangular integral matrices R and S satisfying the shift equivalence equations Definition 5 (i) and (ii). In this case we write $(R, S) : A \sim_{\mathbb{Z}} B$ (lag l).*

Theorem 6 ([1]). *$A \sim B \Longrightarrow A \sim_{\mathbb{Z}} B$. If A, B are primitive, $A \sim_{\mathbb{Z}} B \Longrightarrow A \sim B$.*

Theorem 7 ([1]). *Let A and B be integral matrices, and suppose that R and S are rectangular integral matrices so that $(R, S) : A \sim_{\mathbb{Z}} B$. If $\mu \neq 0$ and $v \neq 0$ are such that $Av = \mu v$, then $w = Sv \neq 0$ and $Bw = \mu w$. Hence A and B have the same set of nonzero eigenvalues, so that $\lambda_A = \lambda_B$.*

Theorem 8 (Perron-Frobenius Theorem, [1]). *Let $A \neq 0$ be an irreducible matrix. Then A has a positive eigenvector v_A with corresponding eigenvalue $\lambda_A > 0$ that is both geometrically and algebraically simple. If μ is another eigenvalue for A, the $|\mu| < \lambda_A$. Any positive eigenvector for A is a positive multiple of v_A. We call λ_A the Perron eigenvalue of A, and v_A a Perron eigenvector of A.*

Theorem 9 (Embedding Theorem, [1]). *Let X and Y be irreducible shifts of finite type. Then there is a proper embedding of X into Y if and only if $h(X) < h(Y)$ and $P(X) \hookrightarrow P(Y)$.*

Theorem 10. *([1]) Suppose that X_G embeds into X_H. Then there are graphs G' and H' so that $X_{G'} \cong X_G$, $X_{H'} \cong X_H$, and G' is a subgraph of H'.*

Note that the Masking Lemma, state below, strengthens this result by showing that we only need to modify H, leaving G exactly as it is, and that G can be realized as a particular kind of subgraph of H'.

Theorem 11 (Masking Lemma, [1]). *Let G and H be graphs. Suppose that X_G embeds into X_H. Then there is a graph K such that $X_K \cong X_H$ and G is an induced subgraph of K.*

References

[1] Douglas Lind and Brian Marcus, *An Introduaction to Symbolic Dynamics and Coding*, First ed., Cambridge University Press., 1995.

[2] Robinson, C., *Dynamical systems: Stability, symbolic dynamics, and chaos*, CRC Press (1995).

[3] Joseph D. Skufca, and Erik M. Bollt, *A concept of homeomorphic defect for defining mostly conjugate dynamical systems*, Chaos, 18: 013008, 2008.

[4] M. Morse and G. A. Hedlund, *Symbolic dynamics*, Amer. J. Math, 60 (1938), 815-866.

[5] Shane Celis, *Partition Matters and Entropy Measures with Logistic Map*, (2009).

[6] Yossi Rubner, Carlo Tomasi and Leonidas. J. Guibas, *A Metric for Distributions with Applications to Image Databases*, IEEE International Conference, (1998).

[7] L. C. Evans, W. Gangbo, *Differential Equations Methods for the Monge-Kantorovich Mass Transfer-Problem*, American Mathematical Society, 1999.

[8] C. Villani, *Topics in Optimal Transportation*, American Mathematical Society, 2003.

[9] Michael Muskulus and Sjoerd Verduyn-Lunel, *Wasserstein distances in the analysis of time series and dynamical systems*, Physica D, 240(2011),45-58.

[10] L. Billings and E. M. Bollt, *Probability density functions of some skew tent maps*, Chaos, Solitions and Fractals, 12 (2001), 365-376.

[11] Hiroshi Teramoto and Tamiki Komatsuzaki, *How does a choice of Markov partition affect the resultant symbolic dynamics?*, Chaos, Solitions and Fractals, 20, 037113 (2010).

[12] Jiongxuan Zheng, Erik M Bollt and Joseph D Skufca, *Regularity of commuter functions for homeomorphic defect measure in dynamical systems model comparison*, DCDS-A Discrete and Continuous Dynamical Systems (2010).

[13] N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Springer, (2002).

[14] A. Lasota and M. C.Mackey, *Chaos, fractals, and noise, Stochastic Aspects of Dynamics*, Springer, 1994.

[15] S. M. Ulam, *Problems in Modern Mathematics*, Wiley, 1970.

[16] B. Marcus and S. Tuncel, *The weight-per-symbol polytope and scaffolds of invariants associated with Markov chains*, Erg. Th. and Dyn. Syst. 11 (1991), 129-180 (2010).

[17] Erik M. Bollt and Naratip Santitissadeekorn, *Applied and Computational Measurable Dynamics*, submitted to SIAM, 2012.