Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Comparison of Systems using Diffusion Maps

Umesh Vaidya, Gregory Hagen, Stéphane Lafon, Andrzej Banaszuk, Igor Mezić, Ronald R. Coifman

Abstract— **In this paper we propose an efficient method of comparing data sets obtained from either an experiment or simulation of dynamical system model for the purpose of model validation. The proposed approach is based on comparing the intrinsic geometry and the associated dynamics linked with the data sets, and requires no a priori knowledge of the qualitative behavior or the dimension of the phase space. The approach to data analysis is based on constructing a diffusion map defined on the graph of the data set as established in the work of Coifman, Lafon, et al.[5], [10]. Low dimensional embedding is done via a singular value decomposition of the approximate diffusion map. We propose some simple metrics constructed from the eigenvectors of the diffusion map that describe the geometric and spectral properties of the data. The approach is illustrated by comparison of candidate models to data of a combustion experiment that shows limit-cycling acoustic oscillations.**

I. Introduction

This paper is concerned with the comparison of dynamical systems data for the purpose of model validation. In many cases, the minimum number of independent variables that are required to describe the approximate behavior of the underlying dynamical system is often much smaller than the dimensionality of the data itself. This is indeed true in many cases of data obtained from fluid flows (see e.g. [6]) and dimensionality reduction in dynamical systems by proper orthogonal decomposition [7] has been extensively studied. Other methods of dimensionality reduction of data include principal component analysis, or PCA, (see e.g. [9]), kernel PCA [17], PCA with multidimensional scaling [18], Laplacian eigenmaps [3], and locally linear embedding [16].

In this work we present an efficient means of comparing dynamical systems data for the purpose of model validation. This work on comparing dynamical system data is inspired by [14]. In this work, a formalism for comparing asymptotic dynamics resulting from different dynamical system models is provided. The formalism is based on the spectral properties of the Koopman operator. A similar approach based on Frobenius-Perron operator (adjoint to Koopman operator [12]) formulism is developed in [13]. In [14], the idea is to construct harmonic averages (in addition to time averages) and use these to analyze the spectral characteristics of the data. The idea in [13] is to use the eigenvectors of Frobenius-Perron operator to analyze the spectral characteristics of the data. These approach captures both the geometry and the dynamics linked with the data set. In [14] [13] the method was applied to the comparison of a noise-driven combustion process with experimental data [8]. The time-series data from both the model and experiment was obtained from a single pressure measurement. The system observable was a collection of indicator functions covering the phase space. The essential ingredients of the metric employed in [14] were a component that captured the spatial, or geometric, properties of the system, and a component that captured the temporal, or spectral, properties of the system.

The theoretical basis that we apply in this paper is established in [5], [10]. In these works, a framework for structural multiscale geometry, organization of graphs, and subsets of $\mathbb{R}^N$ is provided. They have shown that by appropriately selecting the eigenfunctions of the diffusion maps, which describe local transitions, can lead to macroscopic descriptions at different scales. We employ the algorithm developed in [10] to obtain the intrinsic geometry of the data set sampled from the dynamical system.

This paper is organized as follows. In section II, we present the general background of diffusion maps defined on sets of data, as established in [10]. In section III we describe how the eigenfunctions of the diffusion map are used in constructing metrics that capture the spatial and temporal properties of the data sets. In section IV we apply the metrics to compare candidate model data with experimental data. These examples illustrate the necessity of including both spatial and temporal components of the system observable in accurately comparing data.

II. Diffusion maps and diffusion metric

In this section we describe in brief the construction of the diffusion map defined on the data set. This material in this section is from [5],[10]. Given a set of data points Γ construct a weighted function $k(x, y)$ for $x, y \in \Gamma$. $k(x, y)$ is also called as kernel and satisfies the following properties:

- k is symmetric: $k(x, y) = k(y, x)$
- k is positivity preserving: for all x and y in data set X, $k(x, y) \geq 0$
- k is positive semi-definite: for all real-valued bounded functions f defined on X.

$$\int_X \int_X k(x, y) f(x) f(y) d\mu(x) d\mu(y) \geq 0$$

where μ is the probability measure on X. The kernel $k(x, y)$ measures the local connectivity of the data points and

U. Vaidya, G. Hagen, and A. Banaszuk are with United Technologies Research Center, 411 Silver Lane, East Hartford, CT. 06108 VaidyaUG@utrc.utc.com, HagenGS@utrc.utc.com, BanaszA@utrc.utc.com

S. Lafon, research associate, and R.R. Coifman, Professor, are with the Department of Mathematics and Computer Science, Yale University, 51 Prospect St., New Haven, CT 06520. stephane.lafon@yale.edu

I. Mezić, Professor, Department of Mechanical Engineering, University of California Santa Barbara, Santa Barbara CA 93106 mezic@engr.ucsb.edu

0-7803-9568-9/05/$20.00 ©2005 IEEE

hence captures the local geometry of the data points. Several choices for the kernel k are possible all leading to different analyses of data. The idea behind the diffusion map is to construct the global geometry of the data set from the local information contained in the kernel $k(x,y)$. The construction of the diffusion map involves the following steps. First we normalize the kernel $k(x,y)$ in the graph Laplacian fashion [4] . For all $x \in \Gamma$

$$\text{let} \quad v^2(x) = \int_\Gamma k(x,y)d\mu(y)$$

$$\text{set} \quad \tilde{a}(x,y) = \frac{k(x,y)}{v^2(x)}$$

and $\tilde{a}$ satisfies $\int \tilde{a}(x,y)d\mu(y) = 1$. To $\tilde{a}$ we can associate a random walk operator on the data set Γ as

$$\tilde{A}f(x) = \int \tilde{a}(x,y)f(y)d\mu(y).$$

Since we are interested in the spectral properties of the operator it is preferable to work with a symmetric conjugate of $\tilde{A}$. We conjugate $\tilde{a}$ by v in order to obtain a symmetric form and we consider

$$a(x,y) = \frac{k(x,y)}{v(x)v(y)}$$

and operator

$$Af(x) = \int a(x,y)f(y)d\mu(y).$$

The operator A is referred to as diffusion operator. Under very general hypotheses the operator A is compact and self-adjoint so by spectral theory we have

$$a(x,y) = \sum_{j\geq 0} \lambda_j \varphi_j(x)\varphi_j(y), \quad A\varphi_j(x) = \lambda_j\varphi_j(x).$$

Let $a^m(x,y)$ be the kernel of A^m, then at the level of data points the kernel $a^m(x,y)$ has a probabilistic interpretation as a Markov chain with transition matrix a to reach y from x in m steps. The mapping

$$\Phi(x) = (\varphi_0(x), \varphi_1(x), ..., \varphi_p(x), ...)$$

(where φ_i are the eigenfunctions of diffusion operator A) maps the data set $x \in \Gamma$ into the Euclidean space $(\ell^2(N))$, which we will call the diffusion space. Each eigenfunction can be interpreted as a coordinate on the set. This mapping can be used as a diffusion metric to measure the diffusion distance between the data point $x, y \in \Gamma$. More precisely the diffusion metric can be written as

$$D_m^2(x,y) = \sum_{j\geq 0} \lambda_j^m(\varphi_j(x) - \varphi_j(y))^2.$$

For more details on the diffusion metric see [5].

The embedding generated by the eigenfunctions can be used for the dimensionality reduction of the data. For a given accuracy δ we retain only the eigenvalues $\lambda_0, ..., \lambda_{p-1}$ that when raised to the power m, exceed a certain threshold (related to delta) and we use the corresponding eigenfunctions $\varphi_0, \varphi_1, ..., \varphi_{p-1}$ to embed the data points in R^p.

III. Comparison of two data sets

In this section we explain how the theory from the previous section can be used to compare the intrinsic geometry of two data sets. Let us denote by $X = \{x_1, x_2, ..., x_N\}$ and $Y = \{y_1, y_2, ..., y_M\}$ the time series obtained from the experiment and the model simulation respectively. Using time-delayed coordinates we embed the time series data in R^n, where n is sufficiently large. Now we have $N-n$ and $M-n$ data points from experimental time series and model simulation time series, respectively, denoted by

$$\begin{aligned} \bar{X} &:= \{\bar{x}_1, \bar{x}_2, ..., \bar{x}_{N-n}\} \\ \bar{Y} &:= \{\bar{y}_1, \bar{y}_2, ..., \bar{y}_{M-n}\}, \end{aligned} \tag{1}$$

where $\bar{x}_k = (x_k, x_{k+1}, ..., x_{k+n-1})$ and $\bar{y}_k = (y_k, y_{k+1}, .., y_{k+n-1})$. We denote the union of these two data sets by $Z = \{\bar{X}, \bar{Y}\}$. In this paper we use the following Gaussian kernel,

$$k(z_k, z_j) = exp\left(-\frac{\parallel z_k - z_j \parallel^2}{\epsilon}\right), \tag{2}$$

properly normalized in such a way that accounts for the density of the data points [5]. In particular, we are able to handle the case of model and experimental data points with the same geometric structures but different sampling densities. The parameter ϵ specifies the size of the neighborhoods defining the local geometry of the data. The smaller the parameter ϵ the faster the exponential decreases and hence the weight function in (2) becomes numerically insignificant as we move away from the center. It is easy to check that the Gaussian kernel satisfies all the properties of the kernel specified in the previous section.

From this kernel we construct the diffusion operator or the diffusion matrix using the procedure outlined in the previous section. Let $\{\varphi_1, \varphi_2, ..., \varphi_{N+M-2n}\}$ be the eigenvectors of the diffusion matrix and $\{\lambda_1, \lambda_2, ..., \lambda_{N+M-2n}\}$ be the corresponding eigenvalues. Retaining only the first p eigenvectors we can embed the data set Z in a p-dimensional Euclidean diffusion space, where $\{\varphi_1, ..., \varphi_p\}$ are the coordinates of the data points in the Euclidean space. Note that typically $p << n$ and hence we obtain the dimensionality reduction of the original data set. For some index j, the first $N-n$ elements of the eigenvector φ_j are the j-th coordinate in the diffusion space of the $N-n$ data points in X, while the remaining $M-n$ elements are the j-th coordinate in the diffusion space of the data set Y. Denote the eigenvector defined on the experimental data set X by φ^X and the one defined on model data Y by φ^Y. So we have

$$\varphi := \begin{bmatrix} \varphi^X \\ \varphi^Y \end{bmatrix}.$$

Note that the k-th elements of the j-th eigenvectors are given, respectively, by

$$\varphi_{kj}^X := \varphi_j^X(\bar{x}_k), \quad \varphi_{kj}^Y := \varphi_j^Y(\bar{y}_k). \tag{3}$$

A. Comparison of Geometry

We use these eigenvectors to compare the geometry of these two data sets. Various metrics can be used in R^p to compare these two data sets. For convenience, define

$$\phi_k^X = \left(\sum_{j=1}^{p} \lambda_j (\varphi_{kj}^X)^2\right)^{\frac{1}{2}}, \ \phi_k^Y = \left(\sum_{j=1}^{p} \lambda_j (\varphi_{kj}^Y)^2\right)^{\frac{1}{2}} \quad (4)$$

The metrics to be used in this paper are listed below.

1) Weighted average distance

$$D_{avg} = \left[\frac{1}{N-n}\sum_{k=1}^{N-n} \phi_k^X\right] - \left[\frac{1}{M-n}\sum_{k=1}^{M-n} \phi_k^Y\right] \quad (5)$$

2) Pointwise distance
In this metric we compare the distance between the individual data points of the data set in the embedded space (R^p). Hence this metric requires that there are the same number of data points in both the data sets. Let there be N number of data points

$$D_p = \frac{1}{N}\sum_{k=1}^{N} \frac{\left|\phi_k^X - \phi_k^Y\right|}{\phi_k^X}. \quad (6)$$

This metric is sensitive to the ordering of the data set. These are some simple metrics we define for the specific purpose of the example problem in the next section. Of course we can also use a more sophisticated metric like calculating the transportation distance or Hausdorff distance (for details refer to [15]). However the above defined metric gives us satisfactory result for the type of problems we are interested in. The important point is that in the reduced dimensional embedded space the eigenvector of the diffusion map can be used to compare the geometry of the data set.

B. Comparison of Spectra

Since from the very beginning our underlying assumption is that the time series data is sampled from a dynamical system. It is not enough to compare only the geometry of the data set. We also need to compare the temporal properties associated with the geometry of the data set. Here again the eigenvector of the diffusion map can be used to compare the dynamics of the data set.

Information about the dynamics of the data set can be obtained by taking the FFT of the eigenvectors. Denote the FFT of the eigenvectors by

$$\hat{\varphi}_j^X(\omega) := \sum_{k=1}^{N-n} \varphi_{kj}^X e^{-i\omega k}, \ \ \hat{\varphi}_j^Y(\omega) := \sum_{k=1}^{M-n} \varphi_{kj}^Y e^{-i\omega k} \quad (7)$$

To compare the spectral properties of the dynamics we use the L_2 norm for the difference of FFT of the eigenvectors obtained from the experimental and model simulation data, weighted by the corresponding eigenvalues,

$$D_s = \sum_{j=1}^{p} \lambda_j \frac{\left\|\hat{\varphi}_j^X(\cdot) - \hat{\varphi}_j^Y(\cdot)\right\|_2}{\left\|\hat{\varphi}_j^X(\cdot)\right\|_2}. \quad (8)$$

The goal in comparing model simulation data with experimental data is to the identify the parameters of the model so that both the geometry and the spectral properties of the dynamics linked to the geometry of the two data sets are matched.

C. Comparison of Multiple Data Sets

The comparison of data sets is easily extended to the case of data sets from multiple sources. Suppose we have J sets of experimental data $\{X^1 \dots X^J\}$ and K sets of model data $\{Y^1 \dots Y^K\}$. Just as before, we embed each of the time-series data in $\mathbb{R}^n$ resulting in the embedded data sets $\{\bar{X}^1 \dots \bar{X}^J\}$ and $\{\bar{Y}^1 \dots \bar{Y}^K\}$, and take the union $Z = \{\bar{X}^1 \dots \bar{X}^J, \bar{Y}^1 \dots \bar{Y}^K\}$. The eigenvectors can be decomposed accordingly, as

$$\varphi := \left[\varphi^{X^1}, \dots, \varphi^{X^J}, \varphi^{Y^1}, \dots, \varphi^{Y^K}\right]^T.$$

The metrics (5,6,8) that compare any two given sets of data can be applied to the collection of data sets in Z in a straightforward manner.

IV. Simulation results

We consider a simple discrete-time model, as described in [14], to describe a thermo-acoustic system based on the oscillation of an acoustic mode with nonlinear coupling due to the combustion heat-release process,

$$\begin{aligned} x_{k+1}^1 &= (-\alpha + \cos(\omega_0 T_s))\, x_k^1 - \sin(\omega_0 T_s) x_k^1 \\ x_{k+1}^2 &= \sin(\omega_0 T_s) x_k^1 + (-\alpha + \cos(\omega_0 T_s))\, x_k^2 \\ &\quad + K_3 h(K_2 x_{k-N}^1) + K_1 n_k, \end{aligned} \quad (9)$$

where $h(\cdot)$ is some nonlinear function, N is the delay in number of samples, and the model parameters $K_1, K_2, K_3, \alpha, \omega_0$ are determined in order to fit the experimental data. Thermo-acoustic systems are known to exhibit bifurcations between stable and unstable operation (see e.g. [1],[11],[2], among others). When the system is in stable operation, it acts as a noise-driven oscillator with resonance frequency close to ω_0 and amplitude of oscillation depending on a combination of damping α, anti-damping $K_3 h(K_2 x_{k-N}^1)$ and driving noise gain K_1. Under unstable operation, the system exhibits a limit-cycle, with frequency close to ω_0, whose amplitude is determined only from the system damping α and anti-damping effects $K_3 h(K_2 x_{k-N}^1)$. The driving noise acts to perturb the oscillating system from its limit-cycle. Thus, the system is capable of exhibiting two qualitatively different behaviors, which are revealed in terms of the geometry of the phase space. In order to properly validate a model, this geometry must be matched. Furthermore, regardless of whether the system is stable or limit-cycling, such quantitative measures such as pressure oscillation amplitude, frequency, and damping must be matched. The metrics outlined in the previous sections provide an adequate measure of both the qualitative, or geometric, as well as quantitative properties of the system. In this example, the geometrical aspects of the system observable are determined by the amplitude of oscillation of the pressure, the stability

properties of the system, and by the driving noise. The temporal aspects of the system are determined also by the system amplitude and stability, but also by the resonance frequency.

We present three candidate models for comparison to experimental data. In this example, the data was obtained from a single-nozzle combustor rig, as described in [8]. The data shows that the system is operating in a limit-cycle. Each of the three candidate models exhibit different geometric and spectral properties.

Case 1: In first case we compare the experimental time series data with data obtained from the simulation of model (9), where the parameters of the model are chosen such that the system is in stable operation. The qualitative differences in behavior of the two data sets are immediately evident when we examine the histogram of the time series data in Figure 1. The histogram of the experimental data shows a double peak which is indicative of limit-cycling behavior, while the histogram of the model data shows a single peak, indicative of noise-driven stable behavior. Despite this qualitative difference in the data, the amplitude spectra of the raw signals, as shown in Figure 2, are quite similar. Indeed, it is impossible to determine the geometrical properties of the data from the spectra alone. [1]

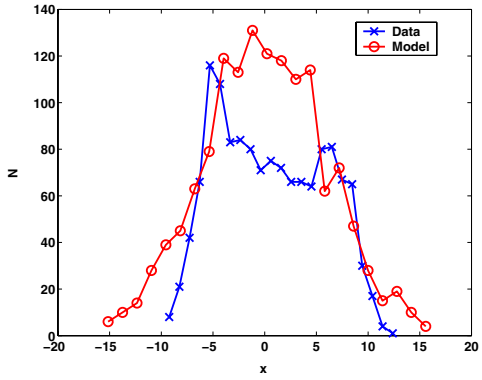

Fig. 1. Histogram of time series data from the data (×) exhibiting limit-cycling behavior and model (○) exhibitting stable behavior.

The eigenvalues of the diffusion map are shown in Figure 3. The zero-th eigenvalue is always unity, corresponding with a flat eigenvector. The next two eigenvalues are greater than the rest of the eigenvalues, indicating that the dimension of the geometry is approximately two. For more details, see [10]. Therefore, the contribution of the projections of the data in to the higher dimensions of the diffusion space appearing in the metrics (4) and (8) are less significant than the projections on to the first two dimensions.

The projections of the data sets on to the first two dimensions of the diffusion space are shown in Figure 4. The limit-cycling data appears to collapse to an approximate circle. The projection of the noise-driven stable data is much more scattered in the diffusion space. Note that the spectra

[1]**For all the figures we follow the following color convention: ×: Experimental data set and ○: Model simulation set**

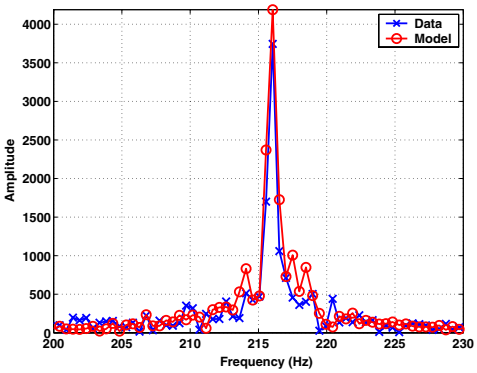

Fig. 2. Amplitude spectra of raw signals from limit-cycling data (×) and noise-driven stable model (○).

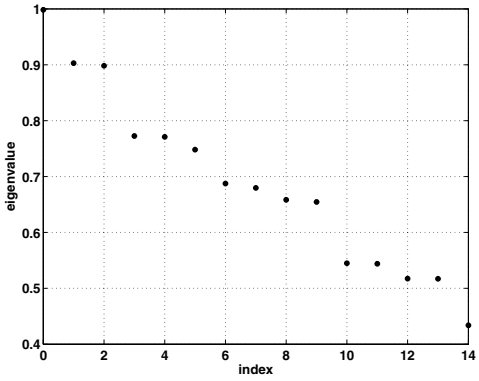

Fig. 3. Eigenvalues of the diffusion map defined on the data sets.

of the eigenfunctions computed by (7) shown in Figure 5 are also nearly identical. The metrics for comparing these two data sets are,

$$D_{avg} = 0.35 \quad D_p = 43 \quad D_s = 0.28$$

This example shows that the spectral comparison metric

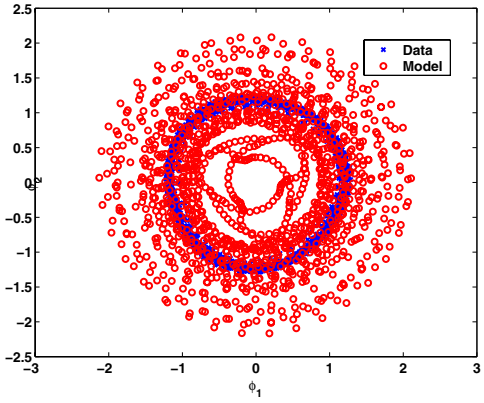

Fig. 4. The projection of the experimental data (×) and model data (○) on to the diffusion space $D_{avg} = 0.35, D_p = 43$.

could show a good match between the data, while their respective geometries differ significantly.

Case 2: In this case, we compare the limit-cycling data of the experiment with a model that also exhibits limit-cycling

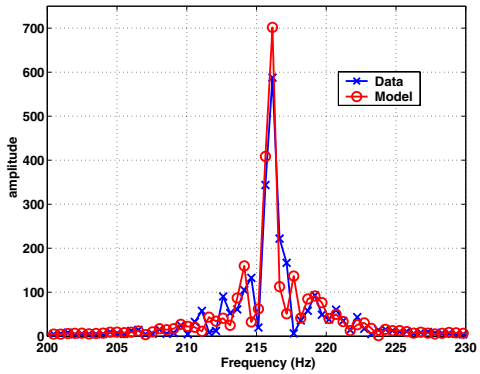

Fig. 5. Amplitude spectra of the projection of the experimental data (×) and model (○) on to the first dimension of the diffusion space. The two spectra are nearly identical $D_s = 0.28$.

behavior, but at a different amplitude and frequency. In this case we expect the geometry of the data sets to be similar, while we expect the spectral comparison to show a mismatch between the data sets.

As in case 1, the eigenvalues of the diffusion map show that both data sets are approximately two-dimensional. The projection of the data on to the two-dimensional diffusion space is shown in Figure 6. Since both model and experiment are limit-cycling, the projections on to the diffusion space look similar. Hence, we would expect the geometry-based metrics to be smaller. The spectra of the projections of the data on to the first dimension of the diffusion space are shown in Figure 7. The spectra are clearly different, showing the necessity of the spectral comparison metric.

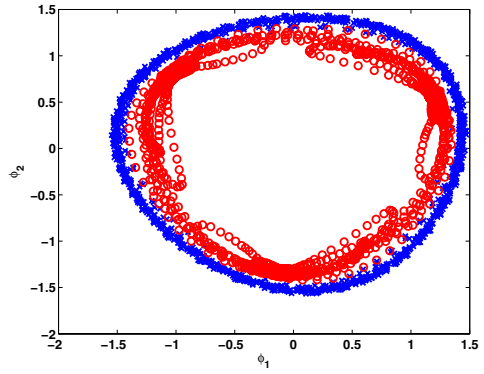

Fig. 6. Projection of experimental data (×) and model data (○) on to the two-dimensional diffusion space $D_{avg} = 0.3267, D_p = 48.90$.

The metrics for comparing these two data sets are,

$$D_{avg} = 0.3267 \quad D_p = 48.90 \quad D_s = 1.6011$$

Note that while the metrics describing the geometry are comparable to case 1, the metric D_s is much larger for this case compared to case 1.

Case 3: In this case the model exhibits a limit-cycle at approximately the same amplitude and frequency as those estimated from experimental data. The projection of the data on to the diffusion space, as shown in Figure 8, along with their spectral content shown in Figure 9 show close agreement. In this case, both the geometric and spectral metrics show a match between the data sets.

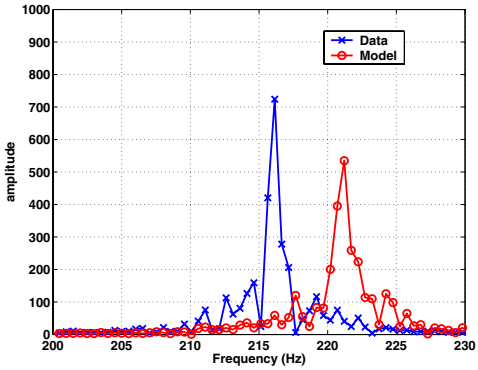

Fig. 7. Amplitude spectra of the projection of the experimental data (×) and model (○) on to the first dimension of the diffusion space. The two spectra are different in amplitude and resonance frequency $D_s = 1.6011$.

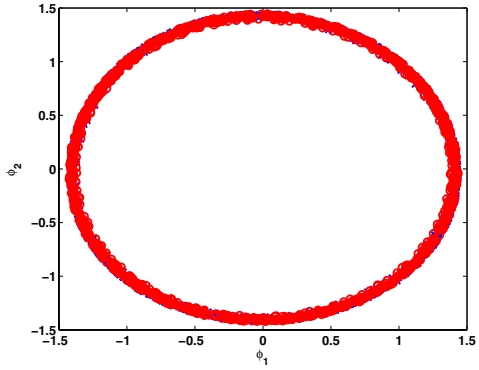

Fig. 8. Projection of experimental data (×) and model data (○) on to the two-dimensional diffusion space. Both sets of data have similar geometry in the diffusion space $D_{avg} = 0.0067, D_p = 35.1506$.

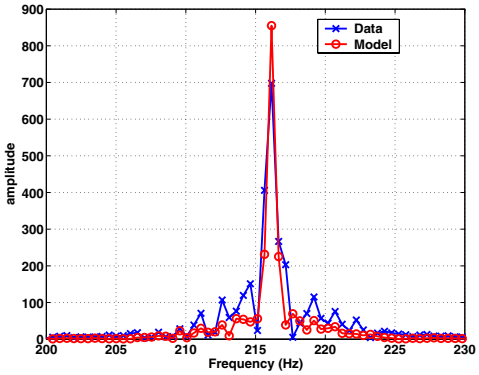

Fig. 9. Amplitude spectra of the projection of the experimental data (×) and model (○) on to the first dimension of the diffusion space. The two spectra are nearly identical $D_s = 0.3969$.

The metrics for comparing these two data sets are,

$$D_{avg} = 0.0067 \quad D_p = 35.1506 \quad D_s = 0.3969$$

Note that the metrics describing the geometry are much

less than those in cases 1 and 2, showing a match in the geoemtery, and the metric D_s is comparable to case 1, showing a match in the spectra.

V. Conclusion

In this paper we have presented a method for comparing dynamical systems data using diffusion maps defined on the data sets. The proposed approach is based on comparing the intrinsic geometry and the associated dynamics linked with the data sets, and requires no a priori knowledge of the qualitative behavior or the dimension of the phase space. We presented some simple metrics constructed from the eigenvectors of the diffusion map that describe the geometric and spectral properties of the data. The approach was illustrated by comparison of three different candidate models to data of a combustion experiment exhibitting limit-cycling acoustic oscillations. Further work will focus on more efficiently combining the metric capturing the geometry and spectral information of the data sets.

Acknowledgements

This work was supported in part by AFOSR contract F49620-01-C-0021, which is gratefully acknowledged.

References

[1] A. Banaszuk, C. A. Jacobson, A. I. Khibnik, and P. Mehta. Linear and nonlinear analysis of controlled combustion processe, part 2: Nonlinear analysis. *Conference on Control Applications, Hawaii, August 1999*, 1999.

[2] M. Basso, R. Genesio, and A. Tesi. A frequency method for predicting limit cycle bifurcations. *Nonlinear Dynamics*, 13:339–360, 1997.

[3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15:1373–1396, 2003.

[4] F. R. K. Chung. *Spectral graph theory*. conference board of the Mathematical Sciences, American Mathematical Society.

[5] R.R. Coifman, S. Lafon, A.B. Lee, M. Maggioni, B. Nadler, F. Warner, and S.W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data, part I: Diffusion maps. *Submitted to Proceedings of the National Academy of Sciences*.

[6] W. R. Graham, J. Peraire, and K. Y. Tang. Optimal control of vortex shedding using low-order models, Part 1: Open-loop model development, and Part 2: Model-based control. *International Journal For Numerical Methods in Engineering*, 44:945–990, 1999.

[7] P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, Cambridge, 1998.

[8] C. A. Jacobson, A. I. Khibnik, A. Banaszuk, J. Cohen, and W. Proscia. Active control of combustion instabilities in gas turbine engines for low emissions. part 1: Phisycs-based and experimentally identified models of combustion instability. *Applied Vehicle Technology Panel Symposium on Active Control Technology, Brauschweig, Germany*, 2000.

[9] I. T. Jollife. *Principal component analysis*. Springer-Verlag, New York, 1986.

[10] S. Lafon. *Diffusion Maps and Geometric Harmonics*. PhD thesis, Yale University, 2004.

[11] T. C. Lieuwen. Experimental investigation of limit-cycle oscillations in an unstable gas-turbine combustor. *Journal of Propulsion and Power*, 18:61–67, 2002.

[12] A. Losata and M. Mackey. *Chaos, Fracrals and Noise: Stochastic Aspects of Dynamics*. Springer-Verlag, New York, 1994.

[13] P.G. Mehta and U.G. Vaidya. On stochastic analysis appraoch for comparing complex system. *Proceeding of 44th conference on Decision and Control and European Control Conference, Spain*, 2005.

[14] I. Mezić and A. Banaszuk. Comparison of systems with complex behavior. *Physica D*, 197:101–133, 2004.

[15] R. Moeckel and B. Murray. Measuring the distance between time series. *Physica D*, 102:187–194, 1997.

[16] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290, 2000.

[17] B. Schölkopf, A. Smola, and K. R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10:1299–1319, 1998.

[18] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290, 2000.