Comparison of Systems with Complex Behavior: Spectral Methods

Igor Mezić[1] and Andrzej Banaszuk[2]
[1]Division of Engineering and Applied Sciences,
Harvard University, Cambridge, MA 02138
[2] United Technologies Research Center, East Hartford, CT 06108.

Abstract— **We present a formalism for comparing the asymptotic dynamics of dynamical systems with physical systems that they model. There is often no need for the detailed (trajectory-wise) comparison of a dynamical system and the physical system that it models, but only comparison in statistical sense. For that purpose, invariant measures are typically considered. But, invariant measures usually can not be observed directly in an experiment. Thus, we base our formalism on time-averages obtained from a single observable. In particular, we constructively prove that, generically, a single observable is needed in order to recover an invariant ergodic measure. Pseudometrics on space of dynamical systems can be defined using this formalism in order to compare their statistical behavior. We also identify the need to go beyond comparing only invariant ergodic measures of systems and introduce an ergodic-theoretic treatment of a class of spectral functionals that allow for this. The formalism is extended for a class of stochastic systems: discrete Random Dynamical Systems. The ideas introduced in this paper can be used for parameter identification and model validation of driven nonlinear models with complicated behavior. As an illustration we provide an example in which we compare the asymptotic behavior of a combustion system measured experimentally with the asymptotic behavior of the model that is a stochastic control dynamical system.**

I. Introduction

This paper is concerned with the issue of comparison of different dynamical system models of physical systems or models of physical systems with the system itself. There are various ways of comparing the behavior of two dynamical systems. All of them involve defining a metric or convergence (of course, once we define a metric we can define convergence). Within the dynamical systems community, this lead the investigation of the above issue in the direction of defining different topologies on spaces of dynamical systems. The definitions of weak and strong topologies for automorphism groups are given in [6], [9]. These are based on the comparison of the action of dynamical systems on open sets of the phase space, and are effectively requirements that the two systems actions be close everywhere. For example, convergence of a sequence of automorphisms T_i to T in the strong topology means that $\{T_i\}$ and T coincide on a larger and larger portion of the phase space as i increases. In the context of modeling, the requirement that the action of two dynamical systems be close everywhere is too strong. Consider, for example systems treated in statistical mechanics. The experimentally established relationship $PV = RT$ is recovered by employing a model consisting of noninteracting particles in a container. This model is certainly not very close to the real dynamics of molecules of monatomic gases for some regimes in the sense that their dynamical actions on the phase space are close. But, they possess the same *time-averaged properties.* Another situation of interest occurs when systems with (formally) infinite number of degrees of freedom are truncated using e.g. Galerkin method to obtain a finite system of ordinary differential equations. In this case, only a proper subset of the initial conditions available for the infinite-dimensional system can be propagated in time by the finite-dimensional truncation, and the comparison in the detailed sense of strong or weak topologies is not possible. These considerations naturally lead to the study of asymptotic dynamics of selected trajectories and this approach was taken in [3], where the stress is on comparing invariant measures. In the case when one of the two systems has a smaller space of initial conditions than the other (e.g. Galerkin projections), projection of invariant measures is used.

While in numerical experiments and analytical work the full state of a system is an observable, in experiments this is typically not the case. Usually the value of one observable - a function on the phase space - is measured. This observation lead to development of the Takens embedding theorem [13], [1]. Here we develop a similar approach using time averages of functions. In particular, we *constructively* prove that ergodic partitions and invariant measures of systems can be compared using a single observable. We develop pseudometrics on spaces of dynamical systems using this result.

In some contexts though, comparing invariant measures is not enough. Consider, for example two systems that have a (geometrically) identical globally attracting limit cycle, but on the limit cycle of the first system the dynamics is given by $\dot{\theta} = \omega_1$ and on the limit cycle of the second system the dynamics is given by $\dot{\theta} = \omega_2$, where $\omega_1 \neq \omega_2$. While these two systems have identical invariant measures supported on the same geometrical object, their *asymptotic speed* is different. We propose here a formalism based on harmonic analysis that extends the concept of comparing the invariant measures. We show that information beyond that obtained using time averages can be acquired by taking harmonic averages if the system has a factor that is a rotation on the circle.

Both the concepts of invariant measure and the harmonic average formalism developed here are related to spectral properties (in particular, the point spectrum) of the so-called Koopman operator U, a linear operator that acts on functions on the phase-space [7]. We stress that in this context questions of identification or validation of asymptotic properties of nonlinear finite-dimensional systems with complex dynamics is transferred to questions of identification or validation of a linear, albeit infinite-dimensional Koopman operator. Here we analyze only the point spectrum of that operator. Our hope is that some of the methods developed in control theory of linear systems can be used to study these issues further.

II. Comparison of long-term dynamics: ergodic partitions and invariant measures

A. *Invariant measures from a single variable*

We are going to consider a dynamical system in discrete time defined by

$$\begin{aligned} x_{i+1} &= T(x_i), \\ y_i &= f(x_i), \end{aligned} \tag{1}$$

where $i \in \mathbb{Z}$, $x_i \in M$, $T : M \to M$ measurable and f a smooth real function on a compact Riemannian manifold M endowed with the Borel sigma algebra. Every dynamical system on a compact manifold possesses an invariant measure μ. We call the function f^* the time average of a function f under T if

$$f^*(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$$

almost everywhere (a.e.) with respect to the measure μ on M. The time average f^* is a function of the initial state x. The operator $P_T : L^1 \to L^1$ such that $P_T(f) = f^*$ is called the time-averaging operator. Note that by Birkhoff's pointwise ergodic theorem [9], f^* exists a.e. for every function $f \in L^1(M)$.

A partition ς of M is defined to be a collection of disjoint sets D^{ς}_{α}, where α is some indexing set, such that $\mu(\underset{\alpha}{\cup} D_\alpha) = \mu(M)$ (see [10]). A product $\varsigma \bigvee \lambda$ of two partitions ς, λ is a partition into sets $D^{\varsigma \bigvee \lambda}_{(\alpha,\beta)} = D^{\varsigma}_{\alpha} \cap D^{\lambda}_{\beta}$ i.e. sets that are intersections of elements of the two partitions. For a finite or countable product ζ of partitions ζ_i, we write $\zeta = \bigvee_i \zeta_i$. The key object in our considerations is partition of the phase space into sets on which the time-averages are constant, i.e. into level sets of f^* : let f be a continuous function on M. The family of sets $C_\alpha, \alpha \in \mathbb{R}$ such that $C_\alpha = (f^*)^{-1}(\alpha)$ is a (measurable) partition of M. We denote this partition by ζ_f and call it the *partition induced by* f.

Every partition ζ_f splits the phase space into sets on which the time-average of f is constant. It turns out that for continuous f the measure zero set on which f^* is not defined is independent of f [8] when M is a compact metric space. For our purpose, the key partition associated with a dynamical system T is its *ergodic partition*: partition of the phase space into sets D on which T is ergodic. More precisely, the ergodic partition ζ_e of M under T is a partition into sets D_α such that on each set D_α there exists an ergodic measure $\mu_{D\alpha}$ such that
1. $\mu_{D_\alpha}(D_\alpha) = 1$,
2. For every $f \in L^1(M), f^*(x \in D_\alpha) = \int_{D_\alpha} f d\mu_{D_\alpha}$ a.e. with respect to μ_{D_α} and
3. For any invariant measure μ, and any measurable set B,

$$\mu(B) = \int_M \mu_{D_\alpha(x)}(B) d\mu(x),$$

where $D_\alpha(x)$ is the element of the partition such that $x \in D_\alpha$.

Our goal is to use time averages obtained from a *single observable* to construct the ergodic partition and thus allow for reconstruction of the ergodic partition from experiments.

Theorem II.1: Let M be a compact Riemannian manifold of dimension m. Let $l/2 > |f|$ and $\kappa_i, i \in \mathbb{N}$ a sequence of continuous periodic functions in $C([-l/2, l/2])$ that is complete. Consider a countable set of functions $f_{i_1,\ldots,i_{2m+1}} = \kappa_{i_1}(f) \cdot \kappa_{i_2}(f \circ T) \cdot \ldots \cdot \kappa_{i_{2m+1}}(f \circ T^{2m})$ (where $i_1, i_2, ..., i_{2m+1} \in \mathbb{N}, i_j \neq i_k$ for any $j \neq k$). Then, for $C^r, r \geqslant 1$ pairs (f, T) it is a generic property that the ergodic partition of a dynamical system T on M is

$$\zeta_e = \bigvee_{i_1,\ldots,i_{2m+1}} \zeta_{f_{i_1,\ldots,i_{2m+1}}}.$$

The essence of the above result is the following. When M is a compact Riemannian manifold, we only need to exhibit a dense countable subset of continuous functions in order to find the ergodic partition. Such a subset is going to be provided by compositions of $(2m+1)-$products of complete set of continuous periodic functions on $\mathbb{R}$ of period l with a generic observable f, i.e. we only need to compute the time-averages of functions

$$\kappa_{i_1}(f(x)) \cdot \kappa_{i_2}(f \circ T(x)) \cdot ... \cdot \kappa_{i_{2m+1}}(f \circ T^{2m}(x)).$$

For example, the set of products of functions $\sin(\frac{2\pi}{l} nx), \cos(\frac{2\pi}{l} mx), \frac{1}{2}$ where none of the integers m, n in the product are the same is a complete set in $C(\mathbb{B})$. The same results hold for systems not defined on compact spaces, but whose attractors are compact sets that are not necessarily manifolds. The extension of Takens theorem for this case can be found in [11].

B. *Pseudometrics*

According to the above description, the asymptotic dynamics partitions the phase space into invariant sets. A sequence of numbers f^*_j is associated with each set in the partition. We can base different pseudometrics on spaces of dynamical systems by using the partition. Let μ be a measure on a compact metric space M. We are going to call systems for which f^* exists for every $f \in L^1(M)$ $\mathcal{B}$-regular [4].

Definition II.1: Let T_1 and T_2 be two continuous, $\mathcal{B}$-regular transformations on M. Then

$$d^a(T_1, T_2) = \max_{f \in C(M)} |P_{T_1}(f) - P_{T_2}(f)|, \tag{2}$$

is called the asymptotic distance between T_1 and T_2.

Example II.2: Consider $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ for which all trajectories tend to $S^1 \in \mathbb{R}^2$ but the dynamics of $T_{1,2}$ on S^1 is given by $\theta \to \theta + \alpha_{1,2}$ where both α_1 and α_2 are irrational. The pseudodistance $d^a(T_1, T_2) = 0$. On the other hand, if α_2 is rational, the pseudodistance is nonzero. This topology is "almost discrete" but with the difference that equivalence classes of transformations with the same statistics are formed.

It was shown in [4] that even if a sequence of transformations $\{T_i\}$ converges to T in the strong topology, T does not need to be $\mathcal{B}$-regular. Thus the strong topology does not provide a good setting to compare statistical properties of systems. But even (2) might not be too useful for comparing the statistics: on one hand it does not distinguish between the systems having very different dynamics but equal statistics like in the example (II.2), and on the other hand it distinguishes between the systems described in the following example:

Example II.3: Let $T_1 : \theta^{i+1} = \lambda\theta_1^i$ and $T_2 : \theta_2^{i+1} = \lambda\theta_2^i + \epsilon$ where $\theta_1, \theta_2 \in S^1$ and $\lambda < 1$. Also, let $f_0(x) = 1/2, f_{2n-1}(x) = \sin(nx), f_{2n}(x) = \cos(nx)$ be a complete set of functions in S^1. The first system has an asymptotically stable fixed point at $\theta_1 = 0$ and the second one at $\theta_2 = \epsilon/(1-\lambda)$. The pseudodistance $d^a(T_1, T_2) > a > 0$ for $\epsilon_m = 1/m$ (i.e..e. a is a lower bound on the pseudodistance) as $\sin(nx)$ is equal to zero at zero and $|\sin(n\ \epsilon/(1-\lambda))| = b > a > 0$ for some n and small a.

As shown in the above example, two systems that have attractors that are very close in space can be distant according to d^a. Thus we come to the point where we define a very natural distance between two systems when we are only interested in matching the asymptotic dynamics on some scale: choose a finite number of functions (i.e. introduce a cut-off) on the phase space and compare the statistics on those. The pseudodistance between T_1, T_2 relative to a function $f : M \to \mathbb{R}$ is defined as

$$d_f^a(T_1, T_2) = |P_{T_1}(f) - P_{T_2}(f)|, \tag{3}$$

The most important property of d_f^a that it renders systems that have "close" attractors "close". In example (II.3), making ϵ smaller would make T_2 converge to T_1 in d_f^a for any smooth f. Obviously, the sum of any number of pseudometrics is a pseudometric. In a specific problem, it is typically easy to identify the important f's. In our thermodynamic example from the introduction it can be the energy of the system. In the case of oscillators, it will be the amplitude of oscillation, etc. The pseudometric (3) is still not entirely satisfactory, as it looses all the "timescale" information about the system. For example, all the irrational rotations on the circle are again identified, as in example II.2.

To treat this problem, we need to extend our formalism to include additional spectral information, and we do so in the next section.

III. Comparison of long-term dynamics: harmonic analysis

In the previous section we have introduced the operator $P_T : f \to f^*$. Note that f^* is an eigenfunction corresponding to eigenvalue 1 of the so-called Koopman operator $U : L^1 \to L^1$, which is defined by

$$Uf(x) = f \circ T(x),$$

as f^* is constant on orbits i.e. $Uf^*(x) = f^*(x)$. The operator P_T can be considered as a member of a family of operators P_T^ω,

$$P_T^\omega(f) = f_\omega^* = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} e^{i2\pi j\omega} f(T^j(x)),$$

where $P_T = P_T^0$.

Example III.1: Consider the maps $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ for which all trajectories tend to $S^1 \in \mathbb{R}^2$ but the dynamics of $T_{1,2}$ on S^1 is given by $\theta' = \theta + \alpha_{1,2}$ where both α_1 and α_2 are irrational. The pseudodistance $d_f^a(T_1, T_2) = 0$ for any continuous (or even L^1) function f. In this case $P^\omega_{T_1^t}(e^{i2\pi\theta}) = 0$ for all $\omega \in S^1, \omega \neq \alpha_1$, while $P^\omega_{T_2^t}(e^{i2\pi\theta}) = 0$ for all $\omega \in S^1, \omega \neq \alpha_2$.

Like the time-averages, the functions f_ω^* also play an important role in the spectral analysis of U: they are the eigenfunctions associated with eigenvalues $e^{-i2\pi\omega}$:

$$\begin{aligned} Uf_\omega^*(x) &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} e^{i2\pi j\omega} f(T^{j+1}(x)) \\ &= e^{-i2\pi\omega} \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} e^{i2\pi(j+1)\omega} f(T^{j+1}(x)) \\ &= e^{-i2\pi\omega} f_\omega^*(x). \end{aligned}$$

That the averages required in the definition of P_T^ω exist almost everywhere was proven by Wiener and Wintner for measure-preserving systems [14]. It is easy to deduce using methods in [14] that this is true for all $\mathcal{B}$-regular T's, as the existence of harmonic averages depends only on the existence of certain autocorrelations which in turn depends on the existence of time-averages of functions. P_T^ω is nonzero only on a countable set of ω's (Lemma in section 4 of [14]). But, when it is non-zero, it can provide substantial new information about the process that we are studying.

It is easy to show that eigenfunctions of U can only be of the form f_ω^*: in fact a nonzero P_T^ω is the orthogonal projection operator onto the eigenspace of U associated with the eigenvalue $e^{-i2\pi\omega}$.

A. Harmonic analysis and ergodic partitions

We would like to know when does the analysis of $P_T^\omega(f)$ bring in new statistical information about T over what we already know using the ergodic partition method. This can be answered using the product $\Pi : M \times S^1 \to M \times S^1, \Pi(x, \theta) = (T(x), \Theta(\theta))$ of T and a rotation $\Theta : S^1 \to S^1$

given by $\theta \to \theta+\omega$. Consider functions $h(x,\theta)=e^{i2\pi\theta}f(x): M\times S^1 \to \mathbb{C}$. As $\theta_j=\theta+j\omega$,

$$h^*(x,\theta)=\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}e^{i2\pi(\theta+j\omega)}f(T^j(x))=e^{i2\pi\theta}P_T^{\omega}(f),$$

and the issue of the properties of $P_T^{\omega}(f)$ is equivalent to the properties of time averages under Π of functions on $M\times S^1$.

The study of Π leads to the following more general question: given two transformations $T_1: M_1 \to M_1$ and $T_2: M_2 \to M_2$, how is the ergodic partition of $T_1\times T_2: M_1\times M_2 \to M_1\times M_2$ related to ergodic partitions of T_1, T_2? Assume that A, B are sets in ergodic partitions of T_1, T_2, respectively. It is certainly not true that $T_1\times T_2$ is ergodic on $A\times B$: consider, for example the cyclic group $C_2=\{-1,1\}$ under multiplication with the measure assigning weight $1/2$ to each element. While translation by -1 is ergodic on C_2, translation by $(-1,-1)$ is not ergodic on $C_2\times C_2$. On the other hand, it is easy to show that no set in $M_1\times M_2$ that is not a subset of some $A\times B$, where A, B are sets in ergodic partitions of T_1, T_2, respectively, is an element of the ergodic partition of $T_1\times T_2$:

Lemma III.2: Let $C\subset M_1\times M_2$ be such that $C\subsetneq A\times B$ where A, B are sets in ergodic partitions of T_1, T_2, respectively. Then $C\notin \zeta_e^{T_1\times T_2}$.

Thus, $A\times B$ is a superset for an element of the ergodic partition of $T_1\times T_2$. It is an element of the partition if $T_1\times T_2$ is ergodic on it. We now return to our original question of the ergodic partition of Π, or rather the question of when $P_T^{\omega}(f)$ bring us statistical information over that provided by the time-averages. Because of the previous lemma, we know that we can restrict attention to a set A on which T is ergodic. Specifically, we ask whether there is a partition of A that contains new statistical information associated with $P_T^{\omega}(f)$. It turns out this is the case when there are eigenfunctions of U that are associated with complex eigenvalues. In the following we relate these eigenfunctions with rotating factors of the map T. Recall that existence of a factor $S: B\to B$ of T is established by proving that there is a measurable *factor map* (or homomorphism) $F: A\to B$ such that $F\circ T=S\circ F$ a.e. and $\mu(F^{-1}(E))=\nu(E)$ for all measurable E, and measures μ,ν, where T preserves μ and S preserves ν. We have the following

Proposition III.3: Let $h_\omega: A\to\mathbb{C}$ be a non-constant eigenfunction of U associated with the eigenvalue $e^{-i2\pi\omega}$. Then h_ω is a factor map and T has a factor that is a rotation on a circle with frequency ω. Conversely, if T admits a factor map to rotation on the circle by angle ω then there is an eigenfunction of T associated with eigenvalue $e^{-i2\pi\omega}$.

Corollary III.4: Let A be a set in the ergodic partition of T. $P_T^{\omega}(f)$ is not constant on A for every $f: M\to\mathbb{C}$ if and only if T has a factor that is a rotation on the circle by an angle $2\pi\omega$.

IV. Stochastic systems

The above theory can be extended to stochastic systems. We will present the application of the above ideas to a stochastic system in the next section and we provide the theoretical framework and relevant results here.

A. Introduction and set-up

For our purpose, the most convenient context in which to analyze stochastic systems is that of Random Dynamical Systems (RDS) [2]. We will work with the Discrete Random Dynamical System (DRDS)

$$\begin{aligned} x_{i+1}&=T(x_i,\xi_i),\\ \xi_{i+1}&=S(\xi_i),\\ y_i&=f(x_i)\end{aligned} \tag{4}$$

where $i\in\mathbb{Z}, x\in M$ a compact Riemannian manifold, $\xi=\{...,\xi^{-1},\xi^0,\xi^1,...\}\in N^{\mathbb{Z}}$, i.e. $\xi^j\in N$, where N is a compact Riemannian manifold endowed with a probability measure p that is absolutely continuous with respect to the Lebesgue measure on N. The product space $N^{\mathbb{Z}}$ is endowed with the standard product measure Ω. S is the shift transformation $S\{...,\xi^{-1},\xi^0,\xi^1,...\}=\{...,\xi^0,\xi^1,\xi^2,...\}$. We consider observables $f: M\to\mathbb{R}$ or $\mathbb{C}, f\in L^1(M)$. We denote $T_\xi^i(x)=T_{\xi_i}\circ...\circ T_{\xi_0}$ where $T_{\xi_i}(x)=T(x,\xi_i)$. We assume that $T_\xi(x)$ is $C^r, r\geq 1$ in x for every $\xi\in N$. With some abuse of notation, we will call the above DRDS T. A probabilistic measure μ on M endowed with the Borel sigma algebra is invariant for measurable T iff

$$\mathbb{E}[\mu(T^{-1}(B,\xi))]=\mu(B)$$

for every measurable B where $\mathbb{E}[\mu(T^{-1}(B,\xi))]=\int_{N^{\mathbb{Z}}}\mu(T^{-1}(B,\xi))d\Omega(\xi)$. The analogue of the Koopman operator is the *stochastic Koopman operator* $U_s: L^\infty(M)\to L^\infty(M)$ defined by

$$U_s f=\mathbb{E}[f\circ T(x,\xi)],$$

where $\mathbb{E}[f\circ T(x,\xi)]=\int_{N^{\mathbb{Z}}}f\circ T(x,\xi)d\Omega(\xi)$. The *expectation of the time-average of* f under T is given by

$$\mathbb{E}f^*(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}U_s^i f(x). \tag{5}$$

The partition of M into level sets of $\mathbb{E}f^*$ is denoted by ζ_f. An ergodic measure on M is a measure μ such that $\mathbb{E}f^*(x)=\int_M f(x)d\mu(x)$ a.e. on M for every $f\in L^\infty(M)$. The ergodic partition ζ_e of M under T is a partition into sets D_α such that on each set D_α there exists an ergodic measure $\mu_{D\alpha}$ such that

1. $\mu_{D_\alpha}(D_\alpha)=1$,
2. For every $f\in L^1(M), \mathbb{E}f^*(x\in D_\alpha)=\int_{D_\alpha}fd\mu_{D_\alpha}$ a.e. with respect to μ_{D_α} and
3. For any invariant measure μ, and any measurable set B,

$$\mu(B)=\int_M \mu_{D_\alpha(x)}(B)d\mu(x),$$

where $D_\alpha(x)$ is the element of the partition such that $x\in D_\alpha$.

B. Ergodic partitions and invariant measures

To state results equivalent to Theorem II.1 we need to use a stochastic version of the Takens embedding theorem. This has recently been provided in [12] (see e.g. Theorem 7 there). In particular, assume that N is a compact manifold and p absolutely continuous with respect to the Lebesgue measure on N. For generic, $C^r, r \geqslant 1$ (f, T) and almost every $\xi \in N^{\mathbb{Z}}$, the map $e : M \to \mathbb{R}^{2m+1}$ given componentwise by $e(x) = (f(x), f(T_\xi x), f(T_\xi^2 x), ..., f(T_\xi^{2m} x))$ is an embedding and thus $e(M)$ is a compact submanifold of $\mathbb{R}^{2m+1}$. Again, it is then necessarily contained in a sufficiently large box $\mathbb{B}$ of side length $l_\xi > 2\max_{x,\xi}|f \circ T(x, \xi)|$ centered at the origin of $\mathbb{R}^{2m+1}$. We can regard $\mathbb{B}$ as a torus $\mathbb{T}^{2m+1}$, i.e. the embedding e can be regarded as a map $e : M \to \mathbb{T}^{2m+1}$.

Theorem IV.1: Let M be a compact Riemannian manifold of dimension m and N a compact manifold of dimension n endowed with a measure p that is absolutely continuous with respect to the Lebesgue measure on N. Let $\kappa_i, i \in \mathbb{N}$ be a sequence of continuous periodic functions in $C([-l_\xi/2, l_\xi/2])$ that is complete. Consider a countable set of functions $f_{i_1,...,i_{2m+1}} = \kappa_{i_1}(f) \cdot \kappa_{i_2}(f \circ T_\xi) \cdot \ldots \cdot \kappa_{i_{2m+1}}(f \circ T_\xi^{2m})$ (where $i_1, i_2, ..., i_{2m+1} \in \mathbb{N}, i_j \neq i_k$ for any $j \neq k$). Then, for almost every ξ, for $C^r, r \geqslant 1$ pairs (f, T) it is a generic property that the ergodic partition of a $C^r, r \geqslant 1$ DRDS T on M is

$$\zeta_e = V \cap \bigvee_{i_1,...,i_{2m+1}} \zeta_{f_{i_1,...,i_{2m+1}}},$$

where $V = \{x \in M | \mathbb{E}(\mathbb{E}(f^* \circ T(x, \xi)) \neq \mathbb{E}f^*(x))$ is of measure 0 with respect to every invariant measure of T.

In the proof the stochastic version of the Takens embedding theorem proven as Theorem 7 in [12] is used.

C. Harmonic analysis

The family of operators $\mathbb{E}P_T^\omega$,

$$\mathbb{E}P_T^\omega(f) \equiv \mathbb{E}f_\omega^* = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} e^{i2\pi j\omega} U_s^i f(T^j(x)),$$

plays the role analogous to the family P_T^ω in the deterministic case.

The system (4) can be also regarded as a control system (see e.g. [5]). When we consider ξ as a control input, the whole ergodic partition on $M \times N^{\mathbb{Z}}$ becomes an interesting object to study. For the discussion of invariant measures in this direction, see e.g. [2].

Now we turn to practical considerations. The concepts defined above allow us to propose procedures for identification of parameters of complex nonlinear systems. We discuss these methods and apply them to experimental data from a combustion experiment in the next section.

D. Probability histograms

The rigorous results stated above suggest that we should take the time-averages of a complete set of continuous functions to study the properties of invariant measures. In applications such as analysis of experimental data, what is typically available is probability histograms. Here we show that these involve a similar construction to the one provided above, a composition of (discontinuous!) functions $\kappa_j : \mathbb{R} \to \mathbb{R}$ with an observable $g : M \to \mathbb{R}$.

In the context of chaotic dynamical systems the probabilistic approach is often taken and a system is described in terms of a histogram of a specific function g on the phase space. Let b be the bin size for the histogram and $z_j \in \mathbb{R}$, $j \in \mathbb{Z}$ a sequence of numbers such that $z_{j+1} = z_j + b$. By the histogram we mean a step function, constant on every interval $I_j = (z_j - b/2, z_j + b/2)$:

$$H_g^T(I_j, x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \kappa_j \circ g(T^i(x)) = \kappa_j^*(x),$$

where $x \in M$. $H_g^T(I_j, x)$ tells us the proportion of time the time-series spends in the interval I_j. The function κ_j is the characteristic function on the interval $I_j = (z_j - b/2, z_j + b/2)$, i.e. $\kappa_j(u) = 1$ if $z_j - b/2 < u < z_j + b/2$ and zero otherwise. If T is ergodic, H is the same function for almost every initial condition x. A possible pseudometric for ergodic systems would be

$$d(T_1, T_2) = \sum_j [H_g^{T^1}(I_j) - H_g^{T^2}(I_j)]^2,$$

where the sum is over some finite set of j's.

The lesson learned from the rigorous study is that we should take time-averages (i.e. histograms) of products of $(\kappa_j \circ g(T^i(x))$ where $i = 0, ..., 2m$, and include them into the pseudometric. The appropriate experimental procedure would be to 1) Get the data from observable g, 2) Determine the dimensionality m of the system using the appropriate embedding theorem, 3) Formulate the model of the same dimension 4) Formulate the histograms of products of $\kappa_j \circ g(T^i(x))$ for experiment and model 6) Compare these histograms in some metric.

Example IV.2: Probability histograms for correlated processes. As mentioned above, one of the common ways of comparing behavior of two systems with complex behavior is taking histograms for a single observable. To show the dangers of this approach when the dynamics of the system is not completely decorrelated and usefulness of the Theorem II.1 in this context, consider two systems having limit cycle attractors in the delay phase space shown in figure 1. Both of the limit cycles are elliptical, one of them having its major axis aligned with the axis z_1, the other with z_2. The dynamics on both limit cycles is assumed to be symmetric with respect to both z_1 and z_2 axis and thus the probability histogram of f (denoted by $p(f)$ in figure 1) is the same for both systems. However, let κ_+ be the indicator function on the interval $(0, l)$. Then $\kappa_+(f) \cdot \kappa_+(f \circ T)$ is the indicator function for the upper right quadrant of the box of side l shown in the figure 1. Clearly the amount of time that these two systems spend in the upper right quadrant is different and thus the time average of this product function reveals the difference in the invariant measures supported on the limit cycles.

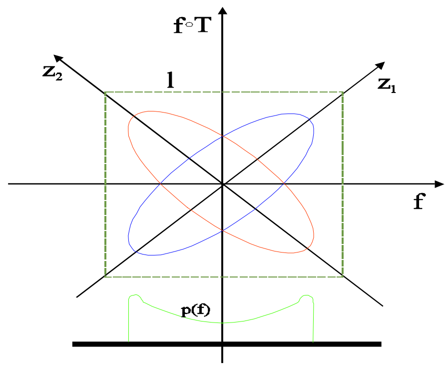

Fig. 1. Two different limit cycles give the same histogram of f.

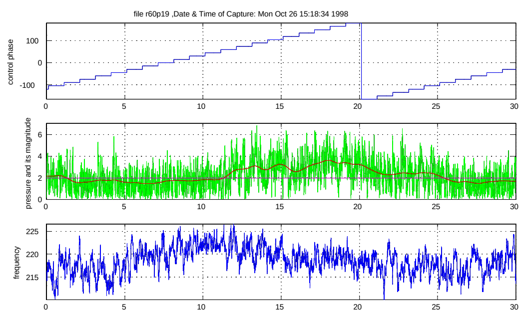

Fig. 2. From the top to the bottom: Control phase, pressure and frequency data. Control phase ramp experiment. Estimates of magnitude and frequency of pressure oscillations are shown. Uncontrolled mean magnitude level is indicated by the solid line in the second figure from top.

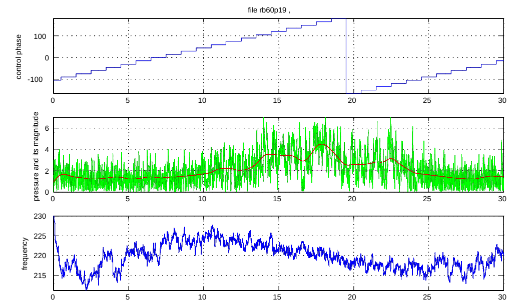

Fig. 3. From the top to the bottom: Control phase, pressure and frequency data. A control phase ramp obtained from a model simualtion.

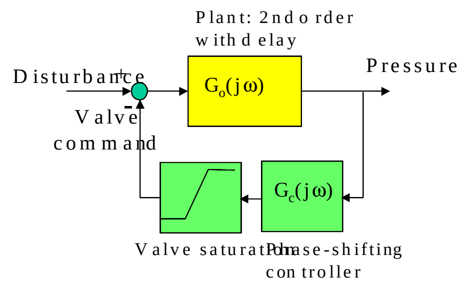

Fig. 4. The block diagram for system model.

V. Analysis of experimental data

In this section we present an example of using a finite number of functionals evaluated on trajectories to identification of parameters of a model describing a United Technologies Research Center combustion rig operating close to an unstable condition associated with a low equivalence ratio. In an experiment a feedback loop has been closed around the combustor with a phase-shifting controller. The control phase has been varied and a time trace of pressure has been recorded. As the phase varies, both attenuation and excitation of the pressure oscillations have been observed. A simple model of the controlled combustion process consists of a linear plant, controller, a saturated actuator, and a driving broad-band noise disturbance. The linear model of the system has been identified from an experimentally obtained frequency response. The model has a form of a lightly damped 2nd order system with delay. An analysis of the linearization of the system at the origin indicates that the linear eigenvalues of the closed-loop system can be moved away or towards the imaginary axis on the complex plane. In particular, it is possible to move the eigenvalues to the right-half plane. In this case, the oscillations would grow and settle on a limit-cycle, as the actuator saturation prevents an unbounded exponential growth. As one adds the strong driving disturbance in the model, an exact analytical study of the behavior of the system is no longer possible. One can still run a numerical simulation and compare the results with experiment. By adjusting the noise level and the value of delay one can qualitatively reproduce the results of the phase ramp experiment in a simulation.

In what follows the delay in the plant and the noise level will be varied to match the results of experiment with closed loop control for several values of control phase. We will define a measure of a good fit by evaluating several long term statistics and combining them with some weights into a one number. For comparison, we also show the phase portraits of the embedded dynamics and find that the parameter values chosen based on pseudometrics ideas result in a good correspondence between the corresponding phase portraits.

For each of the eight values of the control phase from the set (-165, -120, -75, -30, 15, 60, 105, 150) (in degrees) we chose a time interval containing about 0.5 seconds of pressure data (1024 samples of pressure sampled at 2000Hz) corresponding to a constant control phase. Figure 2 shows the experimental results of the control phase ramp experiments where for different control phase values, pressure, mean pressure and frequency are shown. In figure 3 the same data for the model represented in diagram 4 are shown. In figures 5 and 6 we show the time traces, phase portraits, distributions, and PSDs of pressure for control phase $\theta = -30$ and $\theta = 150$ from experiment and from model simulations for delay $\tau = 0.006$ and for two diffrent values of the noise level: $\sigma = 1$ and $\sigma = 3$. The control phases $\theta = -30$ results in a supression of the pressure oscillations. The corresponding closed-loop model is a sta-

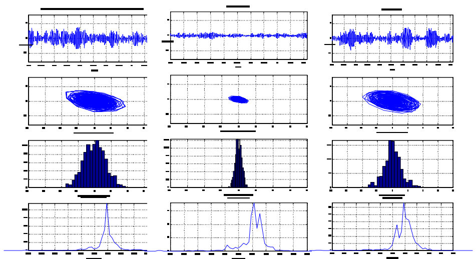

Fig. 5. From the top to the bottom: Time traces, phase portraits, distributions, and PSDs of pressure for control phase $\theta = -30$. Left: from experiment. Middle: from model simulation with $\tau = 0.006$ and $\sigma = 1$. Right: from model simulation with $\tau = 0.006$ and $\sigma = 3$.

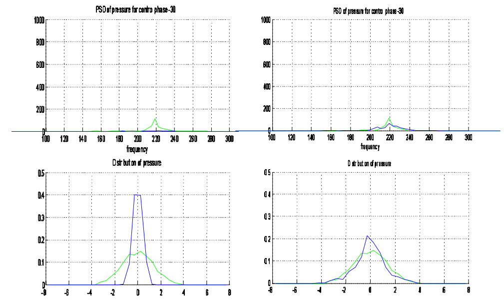

Fig. 7. Comparison pf PSDs and distributions of pressure for control phase $\theta = -30$. Light lines: from experiment. Dark lines: from model simulation.Left: for $\tau = 0.006$ and $\sigma = 1$. Right: for $\tau = 0.006$ and $\sigma = 3$.

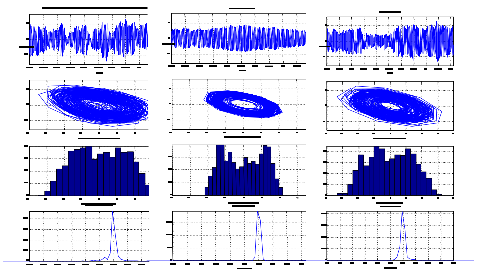

Fig. 6. From the top to the bottom: Time traces, phase portraits, distributions, and PSDs of pressure for control phase $\theta = 150$. Left: from experiment. Middle: from model simulation with $\tau = 0.006$ and $\sigma = 1$. Right: from model simulation with $\tau = 0.006$ and $\sigma = 3$.

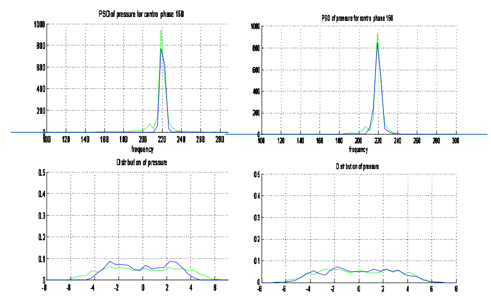

Fig. 8. Comparison pf PSDs and distributions of pressure for control phase $\theta = 150$. Light lines: from experiment. Dark lines: from model simulation.Left: for $\tau = 0.006$ and $\sigma = 1$. Right: for $\tau = 0.006$ and $\sigma = 3$.

ble, noise driven system. On the other hand, the control phase $\theta = 150$ results in an enhancement of oscillations. The model analysis without noise predicts a limit cycling behavior. To visualize the discrepancies between the data from the model simulations and from the experiment, we present the PSDs and distributions of pressure from both models (darker line) on the same plot with the experimental PSDs and distributions (lighter lines) in figures 7 and 8. We can see that the lower value of the noise in the model leads to an underestimation of the pressure oscillation level in experiment for both control phases. We also observe that the distribution of pressure in the data from the high noise model simulation resembles the one from experiment better than the data from the low noise model simulation.

To quantify the discrepancies of the data from experiment and model simulations we would like to choose some quantities that describe the average frequency of the oscillations, the average size of the oscillations, and somehow distinguish a noise-driven limit cycle from a noise-driven stable system. With this goal in mind, for each time interval corresponding to a given control phase θ we calculated the following measures of the long term statistics: 1) The mean frequency of the pressure oscillations. 2) The standard deviation of the frequency pressure oscillations. (**Remark:** Both of these are related to spectral properties discussed in section III.) 3) The standard deviation of the pressure signal. 4) The ratio of pressure measurements in the interval $[-s/2, s/2]$ to the number of all pressure measurement in a given time interval. (**Remark:** 3) and 4) are pseudometrics based on the pressure signal histogram.) In the sequel we will refer to these measures as the *long term statistics*.

A measure of a discrepancy of the statistics from experimental data and model simulation was obtained by taking the absolute value of the difference of the statistics and dividing it by the value of the statistic for the experimental data. A mean value of this relative measure for all eight control phases is called in the sequel a *relative error* for a given statistic. This relative error is a pseudometric in the sense of the theoretical part of this paper.

From the four relative errors we construct a total error measure by adding the four relative errors with the following weights: 1)The weight for the mean frequency of the pressure oscillations: 0.3. 2) The weight for the standard deviation of the frequency pressure oscillations: 0.1. 3) The weight for the standard deviation s of the pressure signal: 0.4. 4) The weight for the ratio of pressure measurements in the interval $[-s/2, s/2]$ to the number of all pressure measurement in a given time interval: 0.2.

In figure 9 we present comparison of the long term statistics from experiment and model simulation for $\tau = 0.006$,

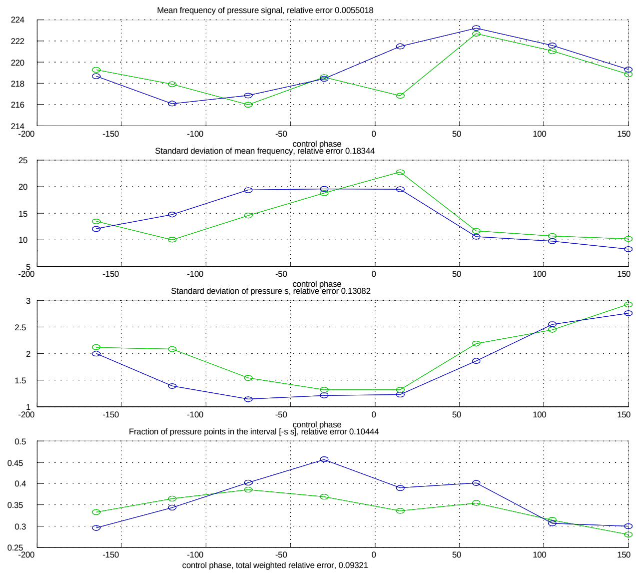

Fig. 9. Long term statistics from experiment and model simulation for $\tau = 0.006$, $\sigma = 3$.

$\sigma = 3$.

The values corresponding to the lowest total relative error of 0.09321 are $\tau = 0.006$ and $\sigma = 3$. The corresponding embedded phase portraits are shown on the right in figures 5 and 6. It is seen that the comparison with experimental phase portrait, shown on the left of the same figures, is quite good.

VI. Conclusions

In this paper we presented some ideas that serve as a framework within which model validation and analysis of nonlinear and/or stochastically driven systems can be done.

Practitioners of experimental and numerical analysis of dynamical systems have found great use of Takens embedding theorem type results. But, embedding methods are often supplemented by statistical considerations such as analysis of probability density functions and spectral analysis. This is specially the case when data is polluted by noise. Here we linked Takens embedding type results with ergodic theory analysis to provide an ergodic-theoretic understanding of probability density and spectral data, both for deterministic and random dynamical systems.

Following the premise that time averages of certain functions on the phase space of a system can be easily obtained experimentally, while complete invariant measures are hard to observe, we have studied the relationship between the two. We have also argued that invariant measures *do not* describe (even in the sense of statistics) everything we would like to know about the asymptotic dynamics of systems. We introduced a family of operators on the space of functions and discussed how the question about the difference of asymptotic dynamics can be transformed into a question on the behavior of this family of operators. Based on this, we introduced pseudometrics on the space of dynamical systems that split this space into equivalence classes of systems having the same (in the sense of the chosen pseudometric) asymptotic dynamics. We presented an example in which this formalism is used to optimize parameters of a model of a combustion experiment.

We stress that questions of identification or validation of asymptotic properties of nonlinear finite-dimensional systems with complex dynamics are in this approach transferred to questions of identification or validation of a linear, albeit infinite-dimensional Koopman operator. Our hope is that some of the methods developed in control theory of linear systems can be used to study these issues further.

On the practical side, we provided a constructive method for obtaining relevant statistics from experiments. This method depends on a *choice* of a particular complete set of periodic functions on an interval. While this choice is irrelevant from the perspective of the theory, as *any* choice of a complete set will give *all* of the required statistical information, the practical issues arising from this are numerous. For example: which complete set do we choose in order to obtain approximate (finite data, finite set of functions) results that are optimal in some sense? We hope to resolve some of these questions in future studies.

VII. Acknowledgements

We would like to thank John Doyle and John Hauser for useful discussions. Andre Valente carefully read the manuscript and suggested various improvements. Igor Mezić was partially supported by an AFOSR grant 98-1-0146, AFOSR grant F49620-98-C-0006 and the Sloan Fellowship. Andrzej Banaszuk was supported by AFOSR grant F49620-98-C-0006.

References

[1] D. Aeyels. Generic observability of differentiable systems. *SIAM Journal on Control and Optimization*, 19:595–603, 1981.
[2] L. Arnold and H. Crauel. Random dynamical systems. *in Springer-Verlag Lecture Notes in Mathematics*, 1486:1–22, 1991.
[3] G. Berkooz. An observation on probability density equations, or, when do simulations reproduce statistics? *Nonlinearity*, 7:313–328, 1994.
[4] D. D'Alessandro, I. Mezić, and M. Dahleh. Some ergodic theorems for sequences of measure preserving transformations. Preprint, 1999.
[5] F.Colonius and W. Kliemann. Lyapunov exponents of control flows. *in Springer-Verlag Lecture Notes in Mathematics*, 1486:1–22, 1991.
[6] P. Halmos. Approximation theories for measure-preserving transformations. *Transactions of the American Mathematical Society*, 55:1–18, 1943.
[7] A. Lasota and M. C. Mackey. *Chaos, Fractals and Noise.* Springer-Verlag, 1994.
[8] R. Mane. *Ergodic Theory and Differentiable Dynamics.* Springer-Verlag, 1987.
[9] K. Petersen. *Ergodic Theory.* Cambridge University Press, Cambridge, 1995.
[10] V. A. Rokhlin. Selected topics from the metric theory of dynamical systems. *Amer. Math. Soc. Transl. Series 2*, 49:171–240, 1966.
[11] T. Sauer, J. A. Yorke, and M. Casdagli. Embedology. *Journal of Statistical Physics*, 65:579–616, 1991.
[12] J. Stark, D. S. Broomhead, M. E. Davies, and J. Huke. Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis*, 30:5303–5314, 1997.
[13] F. Takens. Detecting strange attractors in turbulence. *Springer-Verlag Lecture Notes in Mathematics*, 898:368–381, 1981.
[14] N. Wiener and A. Wintner. Harmonic analysis and ergodic theory. *American Journal of Math.*, 63, 1940.