2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Computation of Uncertainty Distributions in Complex Dynamical Systems

Thordur Runolfsson, Chenxi Lin

Abstract— **The computation of the stationary distribution of an uncertain nonlinear dynamical system is an important tool in analysis of the long term behavior of the system. One common approach is to use a Monte Carlo type method. However, that type of method requires many simulations runs to achieve a reasonable accuracy and can be computationally excessive. In this paper we formulate an alternative approach based on the theory of Random Dynamical Systems to solve this problem. Using the properties of the invariant measure of the Perron-Frobenius operator for the dynamical systems we obtain a simple characterization of the stationary distribution. The state space is discretized to obtain a finite dimensional approximation for the infinite dimensional Perron-Frobenius operator. Furthermore, an efficient subdivision algorithm for state space partition is discussed. The approach is demonstrated through a catalytic reactor system.**

I. Introduction

In this paper we consider an uncertain discrete time dynamical system

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

where the state $x_i \in X \subseteq \mathbb{R}^n$, the uncertainty ξ is a random variable defined on some probability space (Ω, F, P) and taking values in a set $N \subset \mathbb{R}^r$, and y_i is real valued scalar output. The initial state $x_0 = x$ is assumed to be a random variable independent of ξ. Assume that for each fixed ξ and each fixed initial state x the average limit $\overline{y}(x,\xi) = \lim_{i\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} y_i$ exists. Then $\overline{y}(x,\xi)$ is a random variable. The basic problem we are interested in is characterizing the distribution of $\overline{y}(x,\xi)$. We note that if ξ is a fixed parameter and system (1) has the appropriate ergodic properties then the distribution of $\overline{y}(x,\xi)$ is characterized by the stationary or invariant distribution μ_ξ of (1). Similar results are true for the random parameter case as well. Therefore, the characterization of the invariant measure is an important tool in characterizing the long term behavior of the uncertain dynamical system (1). In particular, it gives valuable information about the global dynamics and qualitative behavior of the system.

Obviously, the invariant measure μ_ξ is dependent on the value ξ. Our goal is to compute the invariant measure μ_ξ under the assumption the μ_ξ exists for any ξ. One common approach is a Monte Carlo type method, that samples the distribution of initial state and random parameter and then simulates the system until it reaches stationarity (steady state). Unfortunately, in order to achieve reasonable accuracy one needs many simulation runs and for complex dynamical systems the computational effort may be excessive. Consequently, there is a strong need for efficient alternatives. In this paper we introduce an alternative approach that is based on the fact that the invariant measure is a fixed point of the so-called Perron-Frobenius operator for the dynamical systems.

This research is supported by AFOSR under grant FA9550-05-1-0441.

School of Electrical and Computer Engineering, The University of Oklahoma, Norman, OK 73019, runolfsson@ou.edu, lcx696@ou.edu

II. Mathematical Setup - Random Dynamical Systems

Consider again the dynamical system (1). In order to simplify the discussion we assume that X is a compact subset of $\mathbb{R}^n$ and N is a compact subset of $\mathbb{R}^r$. The random parameter ξ can be very general and is specified in more detail later. We assume that $T(x,\xi)$ is $C^k, k \geq 1$ in x for every $\xi \in N$ and assume that $f : X \to \mathbb{R}$ satisfies $f \in L^1(X)$. Denote $T_\xi^i(x) = T_\xi \circ ... \circ T_\xi$ where $T_\xi(x) = T(x,\xi)$.

Definition 1: Let $\mathcal{M}(X)$ be the vector space of real value measures on X. For a fixed value of ξ the Perron-Frobenius (P-F) operator $P_\xi : \mathcal{M}(X) \to \mathcal{M}(X)$ corresponding to the dynamical system $T_\xi : X \to X$ is defined as

$$P_\xi \mu(B) = \int_{T_\xi^{-1}(B)} d\mu = \mu\left(T_\xi^{-1}(B)\right)$$

for all sets $B \in \mathcal{B}_X$.

We remark that the P-F operator characterizes the evolution of the distribution of the state x_i, i.e. if the initial state has distribution $\nu \in \mathcal{M}(X)$ the distribution of x_i is $P_\xi^i \nu(B)$.

Definition 2: A measure $\mu_\xi \in \mathcal{M}(X)$ is said to be a T_ξ invariant measure if

$$\mu_\xi(B) = P_\xi \mu_\xi(B) = \mu_\xi\left(T_\xi^{-1}(B)\right)$$

for all sets $B \in \mathcal{B}_X$.

Thus, $\mu_\xi \in \mathcal{M}(X)$ is invariant if and only if it is a fixed point of the Perron-Frobenius operator P_ξ.

Let $\mathbf{P} = X \times N$ be the state-uncertainty product space and endow it with the product σ-algebra $\mathcal{P}$ in the usual way, i.e. if $\mathcal{B}_X$ is the Borel $\sigma-$algebra on X and $\mathcal{F}_0$ is the $\sigma-$algebra on N, then $\mathcal{P} = \mathcal{B}_X \times \mathcal{F}_0$.

Definition 3: A probability measure η on $\mathcal{P}$ is called an input measure.

We are interested in the question of how does the uncertainty in the "output" of the process depend on the input measure. For the observable defined by $f : X \to \mathbb{R}$, the "initial" uncertainty is described by a probabilistic measure ϖ_i on $\mathbb{R}$ (endowed with the Borel σ-algebra $\mathcal{B}$) defined by

$$\varpi_i(E) = \eta((f)^{-1}(E)),$$

where $E \in \mathcal{B}$. This measure evolves in time, becoming

$$\begin{aligned}\varpi^n(E) &= \eta((f \circ T_\xi^n)^{-1}(E)) \\ &= \eta((T_\xi^n)^{-1} f^{-1}(E)) = P_\xi^n \eta(f^{-1}(E)),\end{aligned}$$

We call ϖ^n an *output measure*. It describes the uncertainty of the system output at the *n*-th step of the process given the input measure η.

Frequently we are mostly interested in the long term behavior of the solution of the system. In this case the uncertainty in the system output is best studied in terms of the uncertainty in the asymptotic properties of the system. In particular, define the time-average

$$f^*(x,\xi) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T(x_i,\xi)) \tag{2}$$

and the *asymptotic output measure* ϖ_a

$$\varpi_a(E) = \eta((f^*)^{-1}(E)). \tag{3}$$

Next we discuss methods for characterizing the asymptotic output (2) in terms of the invariant measures of the random dynamical system. In particular, we are interested in situations when the system has a physical measure μ in the sense that for almost all $(x,\xi) \in X \times N$

$$f^*(x,\xi) = \int_{X\times N} f d\mu \tag{4}$$

We begin by rewriting system (1) so that it can be studied within the framework of Discrete Random Dynamical System (DRDS) [1]. We are particularly interested in characterizing invariant measures for (1). For the purpose of presenting the formulation in [1] in full generality we allow the uncertainty to vary as a function of time. Let $\Omega = N^{\mathbb{Z}}$ be the space of all sequences taking values in N, denote an element of Ω as ω and let $\mathcal{F}$ be the Borel $\sigma-$algebra on Ω. Let $S(k)$ be the shift transformation on Ω, i.e. $S(k)\omega_i = \omega_{i+k}$, define the map $\pi : \Omega \to N$ by

$$\pi(\omega) = \omega(0)$$

and the coordinate process

$$\xi_i = \omega(i) = \pi(S(i)\omega) = (\pi \circ S(i))(\omega).$$

Let $\mathbb{P}$ be any S invariant (probability) measure. Then $\xi_i = \omega(i)$ is a stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Now we let $\hat{T}(x,\omega) = T(x,\omega(0))$ and get

$$\begin{aligned} x_{i+1} &= T(x_i,\xi_i) = T(x_i,\omega(i)) = \hat{T}(x_i, S(i)\omega) \\ y_i &= f(x_i) \end{aligned} \tag{5}$$

Let $\psi(\omega) : X \to X$ be the operator defined by $\psi(\omega)x = \hat{T}(x,\omega)$. Then the solution of (5) can be represented as $x_i = \varphi(i,\omega)x$ where

$$\varphi(i,\omega) = \begin{cases} \psi(S_{i-1}\omega) \circ \cdots \circ \psi(\omega) & i \geq 1 \\ \mathrm{id}_X & i = 0 \end{cases}$$

where id_X is the identity operator on X. The mapping

$$(\omega, x) \longmapsto (S_i\omega, \varphi(i,\omega)x) = \Theta(i)(\omega, x)$$

is called the skew product of S_i and $\varphi(i,\cdot)$ and is measurable dynamical system on $(\Omega \times X, \mathcal{F} \times \mathcal{B}_X)$. Let π_Ω denote the projection of $\pi_\Omega(\omega, x) = \omega$.

Definition 4: A probability measure μ on $(\Omega \times X, \mathcal{F} \times \mathcal{B}_X)$ is said to be an invariant measure of the DRDS defined by (5) if it satisfies (i) $\Theta(1)\mu = \mu$ and (ii) $\pi_\Omega \mu = \mathbb{P}$.

Note that any measure that is invariant with respect to Θ has a marginal $\pi_\Omega\mu$ that is invariant with respect to S. Furthermore, the second condition in the above definition is imposed since the measure $\mathbb{P}$ is an a-priori specified invariant measure for S . Now let $\mathcal{P}(\Omega \times X)$ be the set of all probability measures on $\Omega \times X$ and define

$$\begin{aligned}\mathcal{P}_{\mathbb{P}}(\Omega \times M) &= \{\nu \in \mathcal{P}(\Omega \times X) : \pi_\Omega \nu = \mathbb{P}\} \\ \mathcal{I}_{\mathbb{P}} &= \{\mu \in \mathcal{P}_{\mathbb{P}}(\Omega \times X) : \mu \text{ is invariant for (5)}\}\end{aligned}$$

Assume $\mu \in \mathcal{P}_{\mathbb{P}}(\Omega \times X)$. A function $\mu_\cdot(\cdot) : \Omega \times \mathcal{B}_X \longrightarrow [0,1]$ is called a factorization of μ with respect to $\mathbb{P}$ if

1) for all $A \in \mathcal{B}_X$, $\mu_\cdot(A)$ is $\mathcal{F}$ measurable,
2) $\mu_\omega(\cdot)$ is a probability measure on $(X, \mathcal{B}_X)$ for $\mathbb{P}$ almost all ω,
3) for all $C \in \mathcal{F} \times \mathcal{B}_X$ we have

$$\mu(C) = \int_\Omega \int_X \chi_C(\omega, x)\, \mu_\omega(dx)\, \mathbb{P}(d\omega)$$

where χ_C is the indicator function for C.

Note that it follows that for $h \in L^1(\mu)$ we have

$$\int_{\Omega\times X} h d\mu = \int_\Omega \int_X h(\omega, x)\, \mu_\omega(dx)\, \mathbb{P}(d\omega)$$

It can be shown that under the assumptions we have made about X the factorization exists and is $\mathbb{P}$ almost surely unique.

Return now back to the uncertain system (1), i.e. assume that for all i we have $\xi_i = \xi$ where ξ is a random parameter that has distribution ρ on N. Then the invariant measure of $S(1)$ is a (random) measure concentrated at ξ and if μ is a physical measure,

$$\begin{aligned} f^*(x,\xi) = \int_{\Omega\times X} f d\mu &= \int_\Omega \int_X f(z)\, \mu_\omega(dz)\, \delta(\omega - \xi)\, d\omega \\ &= \int_X f(z)\, \mu_\xi(dz) \end{aligned}$$

We note that the through the factorization of the physical measure the time average $f^*(x,\xi)$ is an explicit function of the random parameter ξ. The dependence on the (random) initial state can be characterized in terms of the ergodic properties of the system.

Definition 5: Let ϱ be an a-priori fixed measure on X. System (1) is said to be B_ϱ-regular if for each fixed ξ there exists a finite a set of ergodic measures μ_ξ^i, $i = 1, \ldots, K_\xi$ such that for ϱ almost all $x \in M$ and every $g \in C(X)$ there exists a $j \in \{1, \ldots, K_\xi\}$ such that the time average satisfies

$$g^*(x,\xi) = \int_M g(z)\, \mu_\xi^j(dz)$$

Furthermore, there exists an ergodic partition, i.e. disjoint (random) sets $D_1^\xi, \ldots, D_{K_\xi}^\xi$ such that $\varrho\left(M - \cup_{i=1}^{K_\xi} D_i^\xi\right) = 0$ and

$$D_i^\xi = \left\{ x \in M \,\middle|\, g^*(x,\xi) = \int_X g(z)\, \mu_\xi^i(dz) \;\; \forall\; g \in C(X) \right\}$$

The following proposition, that is easily proven using the methods in [2], characterizes the calculation of the asymptotic output measure in terms of the ergodic invariant measure.

Proposition 1: Assume that (1) is parametric B_ϱ*-regular* and the family of measures $\mu_i^\xi, i = 1, ..., K_\xi$ has the property that each $\mu_i^\xi(B)$ is continuous as a function of ξ for any $B \in \mathcal{B}_X$. Assume that the initial measure ν is absolutely continuous with respect to the a-priori measure ϱ. Then the cumulative distribution function F_{ϖ_a} for ϖ_a is piecewise continuous with a finite number of steps.

If we define $\mu_\xi(A) = \sum_{i=1}^{K_\xi} \mu_\xi^i\left(A \cap D_i^\xi\right)$ then $\mu_\xi(A)$ is the invariant measure for T_ξ and factorization of of the invariant measure μ for (1).

The following result proven in [1] provides further insight into the conditions under which the invariant measure of (5) is a physical measure in the sense of (4). The DRDS is said to be continuous if for each fixed ω the mapping $\varphi(\cdot,\omega)\cdot : \mathbb{Z}\times X \to X$ is continuous. We note that by Theorem 1.5.10 in [1] if X is a compact metric space and the DRDS is continuous then $\mathcal{I}_\mathbb{P}$ is a non-empty convex compact subset of $\mathcal{P}_\mathbb{P}(\Omega \times X)$. For a measure $\nu \in \mathcal{P}(\Omega \times X)$ and $f \in L_1(\nu)$ define $\nu(f) = \int f d\nu$.

Theorem 2: Assume that DRDS is continuous. For $\nu \in \mathcal{P}_\mathbb{P}(\Omega \times X)$ define

$$\mu_N(f) = \frac{1}{N}\sum_{n=0}^{N-1} \Theta(n)\,\nu(f) = \frac{1}{N}\sum_{n=0}^{N-1} \nu(f \circ \Theta(n))$$

Then as $N \to \infty$ every limit point of μ_N converges weakly to some $\mu \in \mathcal{I}_\mathbb{P}$ and every $\mu \in \mathcal{I}_\mathbb{P}$ arises in this way.

Finally we have the following general result taken from [1] that further characterizes the factorization of the invariant measure.

Theorem 3: Let $\mu \in \mathcal{P}_\mathbb{P}(\Omega \times X)$. Then (i) $\mu \in \mathcal{I}_\mathbb{P}$ if and only if for all $i \in \mathbb{N}$

$$E\left((\varphi(i,\cdot)\,\mu_\cdot \mid S(i)^{-1}\mathcal{F}\right)_\omega = \mu_{S(i)\omega} \qquad \mathbb{P}\text{–}a.s.$$

(ii) If S is measurably invertible then $S(i)^{-1}\mathcal{F} = \mathcal{F}$ for all $i \in \mathbb{Z}$, and $\mu \in \mathcal{I}_\mathbb{P}$ if and only if for all $i \in \mathbb{Z}$

$$\varphi(i,\omega)\mu_\omega = \mu_{S(i)\omega} \qquad \mathbb{P}\text{–}a.s.$$

Consider now the special case of an uncertain system (1), i.e. $\xi_i = \xi$ for all i where ξ has the distribution ρ on N. In this case, if $\mu \in \mathcal{I}_\mathbb{P}$ then $\varphi(i,\xi)\mu_\xi = \mu_\xi$. Thus, since $\varphi(1,\xi)x = T_\xi(x)$, for each fixed value of $\xi \in N$ the marginal $\mu_\xi = \pi_X \mu$ is the invariant measure of (1) on $(X, \mathcal{B}_X)$. Furthermore, if the dynamical system T_ξ is ergodic for each $\xi \in N$, then the φ-invariant measure μ_ξ is called a random Dirac measure, i.e. there exists a map $h : N \to X$ with $\mu_\xi = \delta_{h(\xi)}$.

Our primary interest is in characterizing the output distribution ϖ_a which in turn implies that we want to characterize the factorization measure μ_ξ for all $\xi \in N$. The most direct way for characterizing μ_ξ is to use Monte Carlo simulation which we discuss next. We then present the main contribution of the paper, i.e. an operator based approach for characterizing μ_ξ that is much more efficient than Monte Carlo based methods for most systems.

III. Monte Carlo Simulation

The most straightforward method for obtaining the stationary distribution (invariant measure) for the uncertain dynamical system (1) is Monte Carlo Simulation. We assume that the probability distributions of the uncertain parameter ξ and the initial state x_0 are known and want to find the output distribution ϖ_a for (1). As we indicated in the previous section this is most easily achieved through the intermediate determination of the invariant distribution μ_ξ for each fixed $\xi \in N$.

We assume the system is B_ϱ-regular and note that in this case the distribution of x_n, characterized by the Perron Frobenius operator $P_\xi^n \nu(A)$, converges to $\mu_\xi(A)$ as $n \to \infty$. Thus $\mu_\xi(A)$ is the distribution of the limiting state.

We propose the following nested Monte Carlo algorithm for calculating the invariant distribution. As before we assume that the parameter ξ has distribution ρ and the initial state has distribution ν and let $\eta = \nu \times \rho$. We sample K points $\xi_1, \ldots, \xi_K$ from the distribution ρ and for each sample ξ we carry out a Monte Carlo simulation for characterizing the distribution $\mu_\xi(A)$. In particular we sample N points x_0^i from the distribution ν and for each of the N samples we simulate the system equation until it reaches steady state $\overline{x}\left(x_0^i, \xi\right)$. Let χ_A be the indicator function for the set $A \subset M$ and define random variables $z_i(\xi) = \chi_A(\overline{x}\left(x_0^i, \xi\right))$. Then compute $\frac{1}{N}\sum_{i=1}^N z_i(\xi)$ as an approximation to $\mu_\xi(A)$.

The accuracy of this method is determined by the number of the samples we pick from the distribution ν. In order to evaluate the error in the Monte Carlo approximation scheme we compare simulation result with the true value $\mu_\xi(A)$. Consider the independent samples $\overline{x}(x_0^i, \xi)$, $i = 1, ..., N$ and assume that the distribution of the $\overline{x}(x_0^i, \xi)$ is identical to that of the steady state $\overline{x}(x, \xi)$, i.e. each simulation had reached the steady state. Then it follows that $E[z_i(\xi)] = E[\chi_A(\overline{x}(x_0^i, \xi))] = \mu_\xi(A)$ and by the central limit theorem

$$e_{KN} = \left(E\left[\left| \frac{1}{N}\sum_{i=1}^N z_i(\xi) - \mu_\xi(A) \right|^2 \right] \right)^{\frac{1}{2}} = \frac{v(\xi)}{\sqrt{N}}$$

where $v(\xi)$ is the standard deviation of the $z_i(\xi)$. We note that the variance depends on the uncertain parameter.

We note that the mean square error between the approximate and true value is proportional to $\frac{1}{\sqrt{N}}$. Therefore, in order to achieve a mean square error of ε we need of the order of $\frac{1}{\epsilon^2}$ simulations, i.e. $N = o\left(\frac{1}{\epsilon^2}\right)$. In many real applications, particularly for complex dynamical systems, each simulation run can be computationally expensive as well

as very time-consuming. Although there exist methods (such as variance reduction schemes) for improving the computational efficiency of Monte Carlo methods as well as modified Monte Carlo methods (such as quasi Monte Carlo method) these methods only result in marginal improvements or are only applicable to a limited class of dynamical systems. Thus, there is a considerable need for efficient alternatives. In the following sections, we introduce a new approach to solve this problem using some recently developed results from Random Dynamical Systems.

IV. Computation of Invariant Measure

In this section we discuss an approach for calculating the invariant distribution for (1) using the theory of random dynamical systems in Section II. In particular we note that under the appropriate conditions the system has an ergodic invariant measure that is characterized as a fixed point of the Perron Frobenius operator P_ξ for T_ξ.

We note that the invariant measure may not be unique. However, it can be shown that, under mild conditions on the dynamical system, by adding small (localized) noise, the resulting system possesses an unique invariant measure [3] that converges to the true ergodic measure as the noise intensity converges to zero. The computational approach relies on the discretization of the P-F operator that we discuss next.

A. Discretization

In order to obtain a finite-dimensional (discrete) approximation of the P-F operator, we consider a finite partition of the state space X, denoted as $B_1, B_2, ...B_k$, where $B_i \cap B_j = \phi$ and $\cup_j B_j = X$. Corresponding to each partition element we associate a positive number $\mu_j \in [0,1]$ with $\sum_{j=1}^k \mu_j = 1$, i.e. $\mu = (\mu_1, ...\mu_k) \in \mathbb{R}^k$ is a probability vector. Define a probability measure on X as

$$\overline{\mu}(dx) = \sum_{i=1}^{k} \overline{\mu}_i \chi_{B_i}(x) \frac{m(dx)}{m(B_i)} \tag{6}$$

where m is the Lebesgue measure and χ_{B_i} is the indicator function for B_i. Then, the action of the Perron Frobenius operator P_ξ on $\overline{\mu}$ on the element B_j is

$$P_\xi \overline{\mu}(B_j) = \overline{\mu}\left(T_\xi^{-1}(B_j)\right) = \sum_{i=1}^{k} \overline{\mu}_i \overline{P}_{ij}(\xi)$$

where the $k \times k$ matrix with entries

$$\overline{P}_{ij}(\xi) = \frac{m(T_\xi^{-1}(B_j) \cap B_i)}{m(B_i)} \tag{7}$$

is a stochastic transition matrix (to simplify notation we drop the ξ dependency of $\overline{P}$). We will see below that the operator $\overline{P}$ is a "good" approximation of P_ξ and the invariant measure for P_ξ can be approximated by a measure $\overline{\mu}$ defined by (6) where the coefficients of $\overline{\mu}$ are invariant for $\overline{P}$, i.e. satisfy $\overline{\mu}_i = (\pi_k)_i$ where $\pi_k = \pi_k \overline{P}$. We note that the computation of the entries of $\overline{P}$ is much more efficient than the Monte Carlo method.

The basic justification for using a finite dimensional approximation for the calculation of the invariant measure lies in the theory of finite dimensional approximations for compact operators [4], [5]. For the Perron Frobenius operator P_ξ we will define an approximate compact operator $P_\varepsilon : L_2(X) \to L_2(X)$ and then use finite dimensional approximations for compact operators to obtain the finite dimensional approximation for P_ξ. Here $L_2(X)$ denotes the space of functions that are square integrable on X. Define a kernel

$$k_\varepsilon(y,z) = \frac{1}{\varepsilon^n m(B)} \chi_B\left(\frac{z-y}{\varepsilon}\right), \qquad x, z \in X$$

where $B \subset X$ is the ball of radius one and center at zero. We note that $k_\varepsilon(T_\xi(x), z)$ is a transition density for the transition function

$$p_\varepsilon(x, A) = \int_A k_\varepsilon(T_\xi(x), z)\, m(dz)$$

It can be shown that $p_\varepsilon(x, \cdot) \to \delta_{T_\xi(x)}(\cdot)$ as $\varepsilon \to 0$ in a weak sense, i.e. p_ε is the transition function for a Markov process that is a small random perturbation of a discrete dynamical system defined by T_ξ [6]. We note that the evolution of the distribution for the Markov process is given by the operator $P_\varepsilon \nu(A) = \int_X p_\varepsilon(x, A)\, \nu(dx)$. If the initial measure ν has density g with respect to m then P_ε can be viewed as an operator mapping $L_2(X) \to L_2(X)$, i.e. the density evolves according to

$$P_\varepsilon g(y) = \int k_\varepsilon(T_\xi(x), y)\, g(x)\, m(dx)$$

Next note that

$$\int \int |k_\varepsilon(T_\xi(x), z)|\, m(dx)\, m(dz) \leq \frac{m(X)}{\varepsilon^n m(B)}$$

Therefore, the transition operator P_ε is a compact operator on $L_2(X)$ [7].

Next we describe how to construct the finite dimensional approximation (7) for the compact operator P_ε that (for small ε) gives a finite dimensional approximation for P_ξ as well [4]. Let V_k be a $k-$dimensional approximation of $L_2(X)$, e.g. $V_k = span\{\varphi_1, \ldots, \varphi_k\}$ for some "independent" functions $\varphi_1, \ldots, \varphi_k \in L_2(X)$. Let $Q_k : L_2(X) \to V_k$ be a projection such that Q_k converges pointwise to the identity in $L_2(X)$ as $k \to \infty$. Define an approximate operator $P_{\varepsilon,k} = Q_k P_\varepsilon$ where P_ε is the compact operator defined previously. Then $\|P_{\varepsilon,k} - P_\varepsilon\|_2 \to 0$ as $k \to \infty$. We use the *finite dimensional* operator $P_{\varepsilon,k}$ as an approximation of the Perron-Frobenius operator P_ξ.

Let V_k be defined by $\varphi_i(y) = \chi_{B_i}(y)$ where the sets B_i form the partition of X discussed earlier. Define the Galerkin projection Q_k of $g \in L_2(X)$ by

$$\langle Q_k g, \varphi_i \rangle = \langle g, \varphi_i \rangle, \quad i = 1, \ldots, k \tag{8}$$

where $\langle \cdot, \cdot \rangle$ is the inner product on $L_2(X)$. Since $\varphi_i(y) = \chi_{B_i}(y)$ we have

$$\int_{B_i} Q_k g = \int_{B_i} g, \quad i = 1, \ldots, k$$

For $g \in V_k = span\{\varphi_1, \ldots, \varphi_k\}$ we write $g(y) = \sum_{i=1}^{k} \varphi_i(y) g_i$. It is easy to see that for any such g we have $\langle g, \varphi_j \rangle = g_j$. Now for any $g \in L_2(X)$ we have by definition $P_{\varepsilon,k} g = Q_k P_\varepsilon g$. Therefore, we get from (8) with $P_\varepsilon g$ replacing g

$$\langle P_\varepsilon g, \varphi_i \rangle = \langle Q_k P_\varepsilon g, \varphi_i \rangle = \langle P_{\varepsilon,k} g, \varphi_i \rangle, \quad i = 1, \ldots, k \quad (9)$$

Since $P_{\varepsilon,k} g \in V_k$ we know that there exist constants $m_i, i = 1, \ldots, k$ such that $P_{\varepsilon,k} g = \sum_{i=1}^{k} \varphi_i(y) m_i$. Furthermore, for $j = 1, \ldots, k$ we get from (9)

$$m_j = \langle P_{\varepsilon,k} g, \varphi_j \rangle = \langle P_\varepsilon g, \varphi_j \rangle$$

Now if in addition $g \in V_k$ then

$$P_{\varepsilon,k} g = P_{\varepsilon,k} \sum_{j=1}^{k} \varphi_j(y) g_j = \sum_{j=1}^{k} P_{\varepsilon,k} \varphi_j(y) g_j$$

Note that if $g = \varphi_j$ then $P_{\varepsilon,k} \varphi_j(y) = \sum_{i=1}^{k} \varphi_i(y) m_{ij}^{\varepsilon}$ where $m_{ij}^{\varepsilon} = \langle P_{\varepsilon,k} \varphi_j, \varphi_i \rangle = \langle Q_k P_\varepsilon \varphi_j, \varphi_i \rangle = \langle P_\varepsilon \varphi_j, \varphi_i \rangle$. Thus

$$P_{\varepsilon,k} g = \sum_{j=1}^{k} \sum_{i=1}^{k} \varphi_i(y) m_{ij}^{\varepsilon} g_j$$

We note that when restricted to the finite dimensional subspace V_k the action of the the operator P_ξ is fully represented by the $k \times k$ matrix M^ε with coefficients m_{ij}^{ε}. We finally note that in the limit $\varepsilon \to 0$ we have $m_{ij}^{\varepsilon} \to m(T_\xi^{-1}(B_i) \cap B_j)$ which after renormalization to a stochastic matrix agrees with (7). Furthermore, it follows from the results in [3], [4] that when the invariant measure of (1) is an ergodic invariant measure in the sense of Definition 5 then the approximate invariant measure π_k converges to μ_ξ as $k \to \infty$.

B. Subdivision

The principal factor affecting the computational complexity in the calculation of the approximate invariant measure is the discretization level on the state space. If the requirement for computational accuracy is not very high we can use a standard subdivision algorithm. As before we assume that the dynamical system is defined on a compact subset of $\mathbb{R}^n$. We start by specifying one box in $\mathbb{R}^n$ that contains X. For a fixed integer L we interactively define 2^L boxes, $B_1, ..., B_{2^L}$, of equal size by a bisection algorithm. The size of the approximate stochastic transition matrix $\overline{P}$ is then $2^L \times 2^L$.We note that if the requirement for computational accuracy is stringent then L will be large and the resulting $\overline{P}$ will be huge requiring excessive storage and computational effort.

The standard subdivision algorithm described above leads to a partition with boxes of equal size. We note that the subdivision is done without utilizing any information about the system dynamics or the invariant measure. However, frequently there exist subsets in the state space that have a very small invariant measure. Consequently, the subdivision of these subsets is not necessary for the computational determination of the invariant measure and their subdivision will lead to unnecessary computation effort. By incorporating information about the invariant measure it is possible to produce more efficient partitioning schemes that result in an adaptive subdivision into boxes of unequal sizes. Here, we introduce an adaptive subdivision algorithm that is a variant of the algorithm developed in [8].

Let $\{\delta_k\}$ be a sequence of positive real numbers such that $\delta_k \to 0$ for $k \to \infty$ and let $\mathcal{B}_k$ be a finite collection of compact subsets of $\mathbb{R}^n$ at step k (the partition at step k). Let π_k be invariant measure stochastic matrix $\overline{P}_k$ obtained at step k. Given an initial pair $(\mathcal{B}_0, \pi_0)$,one inductively obtains $(\mathcal{B}_k, \pi_k)$ from $(\mathcal{B}_{k-1}, \pi_{k-1})$ for $k = 1, 2, \ldots$ in two steps:

1) *Selection and Subdivision:* Define

$$\begin{aligned} \mathcal{B}_{k-1}^{-} &= \{B \in \mathcal{B}_{k-1} : \pi_{k-1}(B) < \delta_{k-1}\} \text{ and} \\ \mathcal{B}_{k-1}^{+} &= \mathcal{B}_{k-1} \backslash \mathcal{B}_{k-1}^{-} \end{aligned}$$

Construct a new collection $\hat{\mathcal{B}}_k^{+}$ such that

$$\bigcup_{B \in \hat{B}_k^{+}} B = \bigcup_{B \in B_{k-1}^{+}} B$$

where

$$\text{diam}(\hat{B}_k^{+}) \leq \theta \ \text{diam}(\hat{B}_{k-1}^{+})$$

for some $0 < \theta < 1$.

2) *Calculation of the invariant measure:* Set

$$\mathcal{B}_k = \mathcal{B}_{k-1}^{-} \cup \hat{\mathcal{B}}_k^{+}$$

For the collection $\mathcal{B}_k$ calculate the invariant measure π_k.

In the realization of the algorithm, we typically subdivide the boxes in the collection $\mathcal{B}_k^{+}$ by bisection and choose the sequence $\{\delta_k\}$ as

$$\delta_k = \frac{1}{N_k} \sum_{B \in \hat{\mathcal{B}}_k} \pi_k(B) = \frac{1}{N_k}$$

where N_k is the number of boxes in $\mathcal{B}_k$.

We illustrate the efficiency of the adaptive subdivision algorithm in a catalytic reactor system in the next section.

V. Example

In [9], [10], a set of nonlinear partial differential equations, which describe heat and mass transfer in a spherically shaped catalytic pellet, was reduced to a dimensionless first-order ordinary differential equation which we represent in the form

$$dx/dt = A - g(x, B)$$

where x is the dimensionless temperature in the reactor, t is time expressed in units of the rise time of the system. The rise time is defined as the time necessary to reach a sufficiently small neighborhood of the equilibrium position. A is the dimensionless external temperature, which is the random parameter in our example. B is a parameter defined by the reacting substances viewed as a constant and $g(x, B)$ is of the form

$$e^{-Bx}(\varphi_1(x)/\varphi_2(x))$$

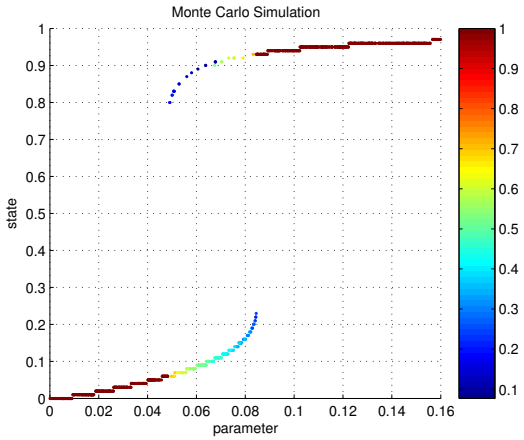

Fig. 1. Invariant measure by Monte Carlo

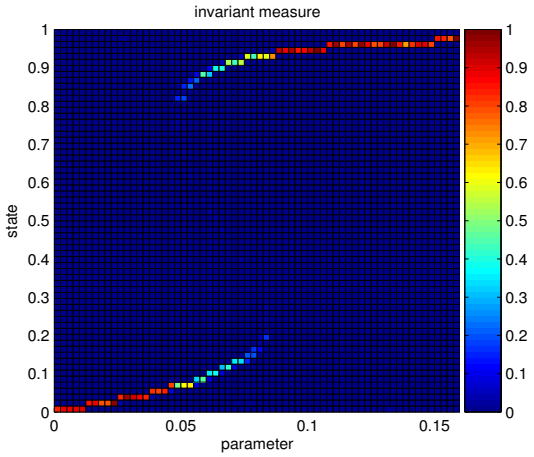

Fig. 2. Invariant measure by standard subdivision

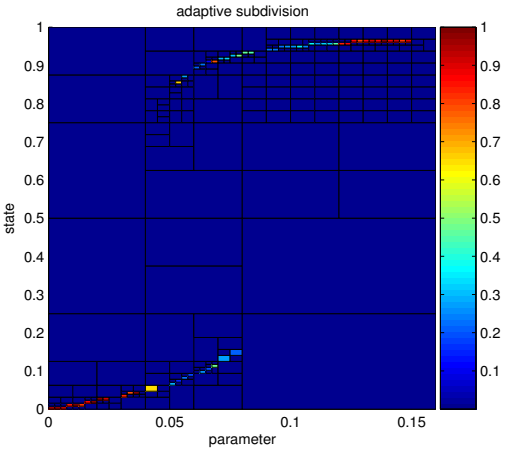

Fig. 3. Invariant measure by adaptive subdivision

where $\varphi_s(x)$, $s = 1, 2$ are polynomials in x. In our simulation, we choose $B = 5.5$ and $g(x, B) = (x/(1-x))e^{-5.5x}$. For initial uncertainty of the system, we assume x has an uniform initial distribution in $[0, 1]$ and A has the same distribution in $[0, 0.16]$. The discretization step is 0.01.

1) Monte Carlo method: We sample 1000 points both in state space and parameter space. As we discuss before, the mean square error is proportional to $\frac{1}{\sqrt{N}}$, that is about 0.03 in this example. The results are shown in Figure 1 and the simulation costs 50.936 sec.

2) P-F method: The standard subdivision is shown in Figure 2 and the adaptive subdivision is shown in Figure 3. For standard subdivision, we have 4096 boxes to cover the whole space with 21.573 sec simulation time while we only have 286 boxes with 8.4858 sec in adaptive subdivision. Moreover, the adaptive subdivision method have much higher accuracy than standard method in computation the invariant measure. The average size of boxes capture the invariant measure is $2.5 * 10^{-6}$ in adaptive subdivision while the size of boxes is $3.9 * 10^{-5}$ in standard subdivision. The adaptive subdivision algorithm shows great efficiency in this example.

VI. Conclusion

In this paper we discussed methods for computing the stationary distribution for dynamical systems with random uncertain parameters. Monte Carlo is the traditional approach for the solution of this problem but it needs many simulation runs in order to achieve reasonable error, which is frequently computationally very expensive as well as time consuming. We introduced an alternative approach using properties of the Perron-Frobenius operator within a unified framework based on measure theoretic concepts from the theory of Random Dynamical Systems. We discretized the state space to obtain a finite-dimensional Markov transition matrix to approximate the infinite-dimensional Perron-Frobenius operator. Two subdivision algorithms to partition the state space were introduced. Finally, we applied both a Monte Carlo method and the P-F based method to a catalytic reactor system for comparison. The simulation result verify the superiority of the P-F method when combined with an adaptive subdivision algorithm for discretization.

References

[1] L. Arnold, *Random Dynamical Systems*. Berlin: Springer, 1998.
[2] I. Mezic and T. Runolfsson, "Uncertainty analysis of complex dynamical systems," in *Proceedings of the 2004 American Control Conference, Boston, MA*, 2004.
[3] Y. Kifer, *Random Perturbations of Dynamical Systems*. New York: Birkhauser, 1988.
[4] M. Dellnitz and O. Junge, "On the approximation of complicated dynamical behavior," *SIAM Journal on Numerical Analysis*, vol. 36, pp. 491–515, 1999.
[5] J. Osborn, "Spectral approximation for compact operators. mathematics of compuation," *Mathematics of Compuation*, vol. 29, pp. 712–725, 1975.
[6] Y. Kifer, "General random perturbations of hyperbolic and expanding trans- formations," *J. Analyse Math*, vol. 29, pp. 47–111, 1986.
[7] K.Yosida, *Functional Analysis*. Springer, 1980.
[8] M. Dellnitz and O. Junge, "An adaptive subdivision technique for the approximation of attractors and invariant measure," *Computing and Visualization in Science*, vol. 1, pp. 63–68, 1998.
[9] C.McGreavy and J.M.Thornton, "Stability studies of single catalyst particles," *Chem.Eng.J.*, vol. 1, pp. 296–301, 1970.
[10] D.S.Cohen and B.J.Matkowsky, "On inhibiting runaway in catalytic recator," *SIAM J.Appl.Math.*, vol. 35, pp. 307–314, 1978.