NUMERICAL METHODS FOR UNCERTAINTY ANALYSIS IN DYNAMICAL SYSTEMS

A Thesis

by

KYUNGEUN KIM

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Swaroop Darbha
Co-Chair of Committee,	Raktim Bhattacharya
Committee Member,	Srinivas Rao Vadali
Head of Department,	Andreas A. Polycarpou

December 2013

Major Subject: Mechanical Engineering

ABSTRACT

The current methods for uncertainty analysis in dynamical systems are restricted in terms of computational cost and evaluation domain since they either use grid points or work only along trajectories. To break through these problems we present a new method: the Rothe & maximum-entropy method which follows the steps below.

A deterministic dynamical system with initial value uncertainties can be analyzed via the uncertainty propagation which is based on the Liouville equation in the form of the first-order linear partial differential equation. On this equation we conduct a semi-discretization in time via A-stable rational approximations of consistency order k and this yields the stationary spatial problem. This spatial problem now can be solved by the spatial discretization scheme: we propose the maximum-entropy approximation which provides unbiased interpolations even with fewer number of scattered points. Through these steps we finally obtain a system of linear equations for the evolution of the probability density function u_t, which can be easily solved in several ways.

This method can provide more efficiency in terms of computation time thanks to using fewer number of scattered points instead of grid points. Also, it enables the constant tracking of probability density functions in a specific fixed domain of interest and this is especially effective for switched systems.

NOMENCLATURE

$\mathcal{P}_t$	Frobenius-Perron operators
$\mathcal{A}$	Infinitesimal generator of the semigroup of Frobenius-Perron operators
$\|\| \cdot \|\|$	$\mathcal{L}_2$-norm
PDF	Probability density function
MOC	Method of characteristics
PDE	Partial differential equation
ODE	Ordinary differential equation

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

1. INTRODUCTION

In real applications, initial values are not known exactly so they can be expressed as statistical uncertainty such as the Gaussian distribution. This uncertainty is propagated forward in time by the system dynamics and the solutions to this uncertainty propagation help understand stochastic systems.

Over the past few years studies on the stochastic system with initial value uncertainties have been done widely [9, 6, 8]; however, the methods to analyze this system are still restricted in terms of computational cost and evaluation domain. One of the methods, Rothe method [20], uses grid points in approximating spatial functions as part of its scheme and this requires a lot of computations at each time step. There are several numerical solutions to reducing computation time in dealing with grid points but for higher dimensional systems the cost becomes more critical even with them. The other method, method of characteristics [8], uses a point cloud and this cloud tends to contract, in most cases of our interest, or expand along trajectories over time depending on the system; as a result, it is not possible to keep evaluating probability density functions in the specific spatial domain of our interest over time.

Thus, our objective is to present a new method to deal with the above two problems: computational cost and evaluation in the specific domain of interest. This can be realized by combining an entropy-based spatial discretization scheme: the maximum-entropy approximation [17] with a temporal semi-discretization scheme: the Rothe method [20].

This thesis is organized as follows. In chapter 2, the concept of uncertainty in dynamical systems and the way of analyzing its time evolution are given to help understand the underlying theories of this thesis. In chapter 3, we introduce the maximum-entropy approximation scheme which is used for the Rothe & maximum-

entropy method, and show two numerical examples using this approximation scheme. In chapter 4, we briefly introduce the current method, the method of characteristics, as a comparison then propose a new method, the Rothe & maximum-entropy method, which first conducts a semi-discretization in time then performs a spatial discretization using the maximum-entropy approximation. At the end of chapter 4, we show two examples using the Rothe & maximum-entropy method, especially for the switched system, it will be shown that this new method can provide effectiveness and accuracy in evaluating probability density functions. And chapter 5 concludes this thesis.

2. UNCERTAINTY IN DYNAMICAL SYSTEMS

In this chapter, we first state the problem to solve throughout this thesis then review some concepts required to understand the effects of uncertainty in deterministic dynamical systems [13]. The uncertainty propagation associated with a dynamical system is defined via time-varying family of transformations of probability distributions called the *semigroup of Frobenius-Perron operators*. These operators describe how a continuous transformation of points in space induces a transformation in the probability distribution over the space.

2.1 Problem Statement

First, let us consider the following autonomous system

$$\dot{\boldsymbol{s}} = \boldsymbol{f}(\boldsymbol{s}, \boldsymbol{p}), \quad \boldsymbol{s}(0) = \boldsymbol{s}_0 \tag{2.1}$$

where $\boldsymbol{s} \in \mathbb{R}^{n_s}$ and $\boldsymbol{p} \in \mathbb{R}^{n_p}$ denote the state and parameters, respectively. The extended state $\boldsymbol{x} := [\boldsymbol{s} \;\; \boldsymbol{p}]^T \in \mathbb{R}^{n_t}$, where $n_t = n_s + n_p$, allows for reconstruction of (2.1).

$$\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 \tag{2.2}$$

The existence and uniqueness of a solution $\boldsymbol{x}(t)$ of this initial value problem is guaranteed by the local Lipschitz continuity of the function $\boldsymbol{F}$ [4], that is,

$$\text{There exists } L \geq 0 \text{ such that } |\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{y})| \leq L\,|\boldsymbol{x} - \boldsymbol{y}|, \quad \forall \boldsymbol{x} \in \mathbb{R}^d,\, \boldsymbol{y} \in B_\kappa(\boldsymbol{x})$$

where $B_\kappa(\boldsymbol{x})$ denotes an open neighborhood of $\boldsymbol{x}$. This property holds throughout the remainder of this thesis.

Next, let $\Phi_t : \mathbb{R}^d \to \mathbb{R}^d$ denote the evolution operator such that $\Phi_t \boldsymbol{x}(0) := \boldsymbol{x}(t)$ then it satisfies the following properties:

- $\Phi_0 \boldsymbol{x} = \boldsymbol{x},\ \forall \boldsymbol{x} \in \mathbb{R}^d$
- $\Phi_t(\Phi_{t'} \boldsymbol{x}) = \Phi_{t+t'} \boldsymbol{x},\ \forall \boldsymbol{x} \in \mathbb{R}^d,\ \forall t, t' \in \mathbb{R}$
- $\Phi_t \boldsymbol{x}$ is differentiable with respect to $\boldsymbol{x},\ \forall t \in \mathbb{R}$

In order to adopt uncertainty to the system, we assume that $\boldsymbol{x}_0 = \boldsymbol{X}_0$ is a random variable. Under this assumption, the evolution $\boldsymbol{X}_t$, defined by $\Phi_t \boldsymbol{x}_0$, is a random variable as well.

Now, we can define the following initial value uncertainty problem [20].

Definition 2.1.1 (Initial Value Uncertainty Problem): For the system $\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x})$, assume that the initial value $\boldsymbol{x}_0 = \boldsymbol{X}_0$ is a random variable and has a known probability distribution with its density u_0. The problem is to compute the probability density function u_t associated with the random state $\boldsymbol{x}(t) = \boldsymbol{X}_t$ on a finite interval $t \in [0, T]$ where $u_t = u(t, \cdot)$, $u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, $t \geq 0$.

2.2 Uncertainty Propagation

The evolution operator Φ_t for a fixed time $t \geq 0$ accounts for a transformation on the state space Ω and this transformation results in a change to the probability distribution on Ω. The probability of a set B at time t should be equal to the probability of its pre-image $\Phi_t^{-1}(B)$ as depicted in Fig. 2.1 and this property enables us to draw the following relation between u_0 and u_t:

$$u_t = \mathcal{P}_t u_0 \tag{2.3}$$

where $\mathcal{P}_t$ corresponds to Φ_t and is called the *Frobenius-Perron operator*. The Frobenius-Perron operator maps an initial probability density u_0 to u_t as Φ_t does an initial state $\boldsymbol{x}_0$ to $\boldsymbol{x}(t)$ at time t. The general definition of the Frobenius-Perron operators is as follows.

Definition 2.2.1 (Frobenius-Perron operator): The *Frobenius-Perron* operator $\mathcal{P}_t : \mathcal{D} \to \mathcal{D}$ (where $\mathcal{D}$ is the space of probability distributions on the manifold M) associated with the diffeomorphism $\varphi_t : M \to M$ is a so-called *transfer operator* on the space of probability distributions, defined by the change-of-variables formula

$$\mathcal{P}_t u = \int_M u(\varphi_t^{-1}(x)) \det \left(D\varphi_t^{-1}(x) \right) \, \mu(dx), \tag{2.4}$$

where $\mu(dx)$ is Lebesgue measure on M and $u \in \mathcal{D}$ is arbitrary. Furthermore, the family of operators $\{\mathcal{P}_t\}_{t \geq 0}$ is a *semigroup of Frobenius-Perron operators* satisfying

- $\mathcal{P}_0 u = u, \quad \forall u$
- $\mathcal{P}_t(\mathcal{P}_s u) = \mathcal{P}_s(\mathcal{P}_t u) = \mathcal{P}_{t+s} u, \quad \forall u, \ t, s \geq 0$

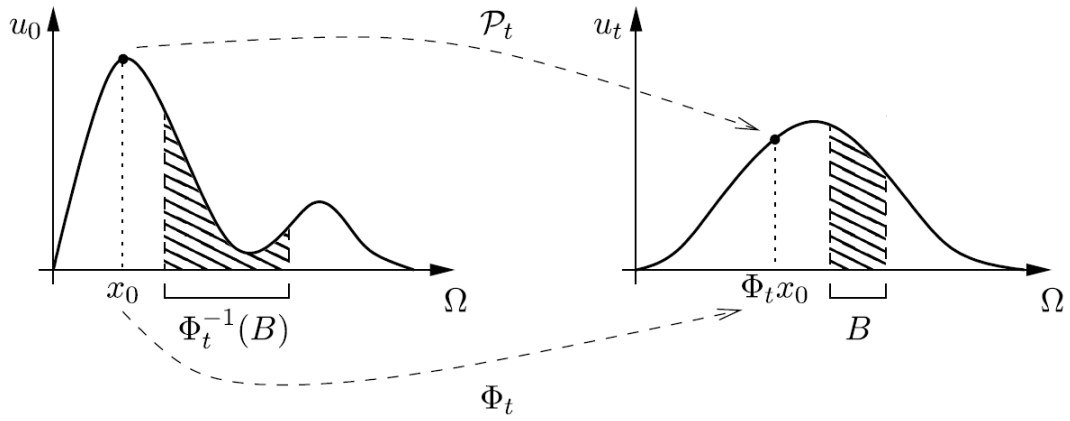

Figure 2.1: The conservation of probability mass on any set B and its pre-image $\Phi_t^{-1}(B)$ defines the Frobenius-Perron operator $\mathcal{P}_t$ corresponding to Φ_t, which relates the two density functions u_0 and u_t to each other. [20]

where a semigroup is a set S together with a binary operation $\cdot$ that satisfies the associative property: $(a \cdot b) \cdot c = a \cdot (b \cdot c),\ \forall a, b, c \in S$.

As defined above, the Frobenius-Perron operator describes how some initial distribution of system uncertainty $u \in \mathcal{D}$ is propagated forward in time by the system dynamics; see [12, 13] for more discussions and examples on the uncertainty propagation. In addition, an infinitesimal change in the system uncertainty is characterized by the following proposition [13].

Proposition 2.2.2: Denote the semigroup of Frobenius-Perron operators associated with a *deterministic* dynamical system by $u_t(x) = u(t, x) = \mathcal{P}_t u_0(x)$, where $u_0 \in \mathcal{D}$ is given. Then the *Liouville equation*

$$\frac{\partial}{\partial t} u_t(x) + \sum_{i=1}^{n} \frac{\partial}{\partial x_i} (u_t(x)\, f(x)) = 0 \tag{2.5}$$

is the unique *infinitesimal generator* of $u_t(x)$.

In this thesis we are only focused on a deterministic dynamical system but it is also notable that the infinitesimal generator of the semigroup of Frobenius-Perron operators for a *stochastic* dynamical system is a partial differential equation called the *Fokker-Planck equation* [13] which describes the propagation of system uncertainty when the system evolution is random.

3. MAXIMUM-ENTROPY APPROXIMATION

In this chapter we introduce the maximum-entropy approximation which is used as a spatial discretization scheme for the Rothe method in the next chapter. The notion of *entropy* in information theory was first introduced by Shannon as a measure of uncertainty [15] and Jaynes further proposed the *maximum-entropy principle* in which maximizing entropy enables the least-biased statistical inference under the condition of insufficient data [10, 11]. Based on these two works, Sukumar proposed the *maximum-entropy approximation* [17, 19] which is used to construct basis functions for the Rothe & maximum-entropy method.

3.1 Construction of Basis Functions

Suppose that a sample space, Ω, consists of mutually independent discrete events $x_1, x_2, \ldots, x_n$ which occur with unknown probabilities $p(x_1), p(x_2), \ldots, p(x_n) \geq 0$, respectively but with a known expected value $E(x)$. Since $P(\Omega) = 1$ for a random experiment, probabilities $p(x_i)$ should satisfy

$$\sum_{i=1}^{n} p(x_i) = 1 \tag{3.1}$$

and since the expected value is known, they should also satisfy

$$\sum_{i=1}^{n} p(x_i)\, x_i = E(x) \tag{3.2}$$

Then the most likely probability distribution $p(x_i)$ can be obtained by maximizing the following Shannon entropy for a discrete probability distribution: [15, 10]

$$H(\boldsymbol{p}) = -\sum_{i=1}^{n} p(x_i)\, \ln(p(x_i)) \tag{3.3}$$

Similarly, consider a convex hull, Ω, consisting of a set of distinct scattered nodes $\{\boldsymbol{x}_i\}_{i=1}^n$ in $\mathbb{R}^d$. Any point $\boldsymbol{x} \in \bar{\Omega}$ where $\bar{\Omega} = \Omega \cup \partial\Omega$ has basis functions $\phi_i(\boldsymbol{x})$ associated with each node $\boldsymbol{x}_i$. Since the basis function $\phi_i(\boldsymbol{x})$ can be regarded as a probability of influence of node $\boldsymbol{x}_i$ on the point $\boldsymbol{x}$, $\phi_i(\boldsymbol{x}) \geq 0$ and we have the following problem from (3.3)

$$\max_{\{\phi_i\}} \left(H(\boldsymbol{\phi}) = -\sum_{i=1}^{n} \phi_i(\boldsymbol{x}) \ln \phi_i(\boldsymbol{x}) \right) \tag{3.4}$$

Also, (3.1) and (3.2) similarly yield the following constraints, respectively:

$$\sum_{i=1}^{n} \phi_i(\boldsymbol{x}) = 1 \quad \text{(Constant precision)} \tag{3.5}$$

$$\sum_{i=1}^{n} \phi_i(\boldsymbol{x})\, \boldsymbol{x}_i = \boldsymbol{x} \quad \text{(Linear precision)} \tag{3.6}$$

The constant and linear precisions are sufficient conditions for convergence in the Galerkin method for second-order partial differential equations [16].

The general form of entropy for a discrete probability distribution is given by [10, 11]

$$H(\boldsymbol{p}, \boldsymbol{m}) = -\sum_{i=1}^{n} p_i \ln \left(\frac{p_i}{m_i} \right) \tag{3.7}$$

where p_i is a probability of event i and m_i is a prior estimate of p_i. This is called the Shannon-Jaynes entropy and yields the following constrained optimization problem:

$$\max_{\boldsymbol{\phi} \in \mathbb{R}_+^d} \left(H(\boldsymbol{\phi}, \boldsymbol{m}) = -\sum_{i=1}^{n} \phi_i(\boldsymbol{x}) \ln \left(\frac{\phi_i(\boldsymbol{x})}{m_i(\boldsymbol{x})} \right) \right) \tag{3.8}$$

subject to the linear constraints:

$$\sum_{i=1}^{n} \phi_i(\boldsymbol{x}) = 1 \tag{3.9}$$

$$\sum_{i=1}^{n} \phi_i(\boldsymbol{x})\boldsymbol{x}_i = \boldsymbol{x} \tag{3.10}$$

For d-dimensional problems, let $\lambda_r (r = 0, 1, \ldots, d)$ be the Lagrange multipliers for the $(d+1)$ linear constraints then we have the following variational equation [10].

$$\begin{aligned}\delta \Bigg[\sum_{i=1}^{n} \left(-\phi_i \ln \frac{\phi_i}{m_i} \right) + \lambda_0 \left(1 - \sum_{i=1}^{n} \phi_i \right) + \lambda_1 \left(x - \sum_{i=1}^{n} \phi_i x_i \right) \\ + \lambda_2 \left(y - \sum_{i=1}^{n} \phi_i y_i \right) + \lambda_3 \left(z - \sum_{i=1}^{n} \phi_i z_i \right) + \cdots \Bigg] = 0\end{aligned} \tag{3.11}$$

where $\boldsymbol{x}_i = (x_i, y_i, z_i, \cdots)^T$ and $\boldsymbol{x} = (x, y, z, \cdots)^T$. By letting $\lambda_0 = \ln Z - 1$ and using (3.9) where Z is known as the partition function and used for the canonical distribution in statistical mechanics [19], the solution of this equation can be written as

$$\phi_i(\boldsymbol{x}) = \frac{Z_i(\boldsymbol{x})}{\sum_{j=1}^{n} Z_j(\boldsymbol{x})} = \frac{Z_i(\boldsymbol{x})}{Z(\boldsymbol{x})} \tag{3.12}$$

where

$$Z_i(\boldsymbol{x}) = m_i(\boldsymbol{x}) \exp(-\boldsymbol{x}_i^T \boldsymbol{\lambda}(\boldsymbol{x})) \tag{3.13}$$

The prior estimate, $m_i(\boldsymbol{x})$, can be selected out of global radial basis functions, compactly supported radial basis functions, cubic spline weight functions, etc. In this thesis we use the global radial basis function: the Gaussian, $m_i(\boldsymbol{x}) = \exp(-\beta \|\boldsymbol{x}_i -$

$\boldsymbol{x}\|^2)$ where β is a constant variable. By letting $\tilde{\boldsymbol{x}}_i = \boldsymbol{x}_i - \boldsymbol{x}$, (3.12) and (3.13) become

$$\phi_i(\tilde{\boldsymbol{x}}) = \frac{\tilde{Z}_i(\tilde{\boldsymbol{x}})}{\sum_{j=1}^{n} \tilde{Z}_j(\tilde{\boldsymbol{x}})} = \frac{\tilde{Z}_i(\tilde{\boldsymbol{x}})}{\tilde{Z}(\tilde{\boldsymbol{x}})} \tag{3.14}$$

$$\tilde{Z}_i(\tilde{\boldsymbol{x}}) = m_i(\tilde{\boldsymbol{x}}) \exp(-\tilde{\boldsymbol{x}}_i^T \boldsymbol{\lambda}(\tilde{\boldsymbol{x}})) \tag{3.15}$$

Solving these equations for $\boldsymbol{\lambda}$ can be recast as the following dual problem for the convergence of solutions [1]:

$$\text{Find } \boldsymbol{\lambda} \text{ such that } F = \ln \tilde{Z}(\boldsymbol{\lambda}) \text{ is minimized.} \tag{3.16}$$

Here F is an upper bound for the entropy H [2] and more details can be found in [18]. This dual problem can be solved through various optimization schemes and we use the MATLAB command *fminsearch* to find its solution in the numerical examples.

Once we find $\boldsymbol{\lambda}$ minimizing F, the maximum-entropy basis functions are obtained from Equation (3.14). Furthermore, given measurements $\{z_i\}_{i=1}^n$ for the nodes $\{\boldsymbol{x}_i\}_{i=1}^n$, the maximum-entropy basis matrix function A is constructed then we can find a linear estimator $\hat{\boldsymbol{\alpha}}$ from the following estimation problem:

$$\text{Given } A\hat{\boldsymbol{\alpha}} = \boldsymbol{b}, \text{ find } \hat{\boldsymbol{\alpha}} \text{ that minimizes } \|A\hat{\boldsymbol{\alpha}} - \boldsymbol{b}\|_2 \tag{3.17}$$

where

$$A = \begin{bmatrix} \phi_1(\boldsymbol{x_1}) & \cdots & \phi_n(\boldsymbol{x_1}) \\ \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{x_n}) & \cdots & \phi_n(\boldsymbol{x_n}) \end{bmatrix} \quad \hat{\boldsymbol{\alpha}} = \begin{bmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_n \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

There are a few algorithms to find $\hat{\boldsymbol{\alpha}}$ such as:

- Least Squares (cheap but less accurate)
- QR Decomposition
- Singular Value Decomposition (expensive but more reliable)

The matrix A computed from the maximum-entropy scheme tends to have low rank when a lot of data points are given as nodes and they are close enough to each other. Unfortunately, all of the three algorithms do not work out well for our problem so we employ the following *Truncated Singular Value Decomposition (SVD)* which gives fast and reliable estimates even with fewer data points.

Definition 3.1.1 (Truncated Singular Value Decomposition): If A is an $m \times n$ matrix and has rank r, $A = U_{m\times m}S_{m\times n}V^T_{n\times n}$ where U and V are unitary matrices, and S is a $m\times n$ diagonal matrix consisting of the singular values of A. By eliminating zero or nearly zero singular values of S and making it a $r \times r$ diagonal matrix, and transforming U and V into $m \times r$ and $n \times r$ matrices, respectively, we obtain a new matrix $A = U_{m\times r}S_{r\times r}V^T_{n\times r}$.

In approximating probability density functions at each time step, it is important to maintain the shape feature of them since even a spike with the overall error kept small may result in a totally different shape of probability density functions at the next time step. This means that having a similar shape with a higher overall error is more desirable than having a spike with a lower overall error. Here the maximum-entropy approximation has a benefit: the maximum-entropy basis functions enable the least-biased statistical interpolations and this means that this approximation is the best tool in extracting the feature of probability density functions in terms of shapes at each time step, although the method of least squares using other basis functions might be a bit better in terms of overall errors.

3.2 Numerical Examples

In this section we apply the maximum-entropy basis functions and the truncated SVD to two systems and demonstrate how well these two methods interpolate them based on some given points.

First, suppose we have the system: $z = (x^2 - 1)(10y^2 - 4) + 4y^2 + 2$. Fig. 3.1 shows true and interpolated z values for 51×51 points based on $\hat{\boldsymbol{\alpha}}$ estimated by the maximum-entropy basis functions of given 10×10 points. The Frobenius norm of the errors for 51×51 points is 1.4908.

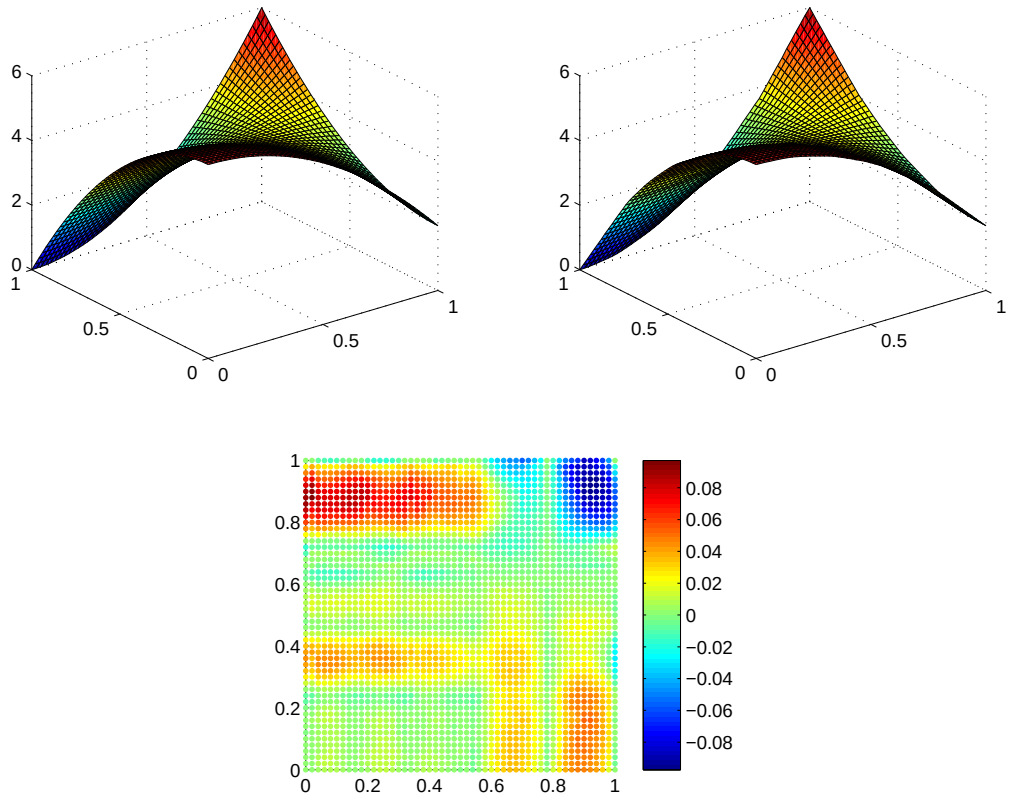

Figure 3.1: Top-left: true data; Top-right: interpolated data; Bottom: errors at each point

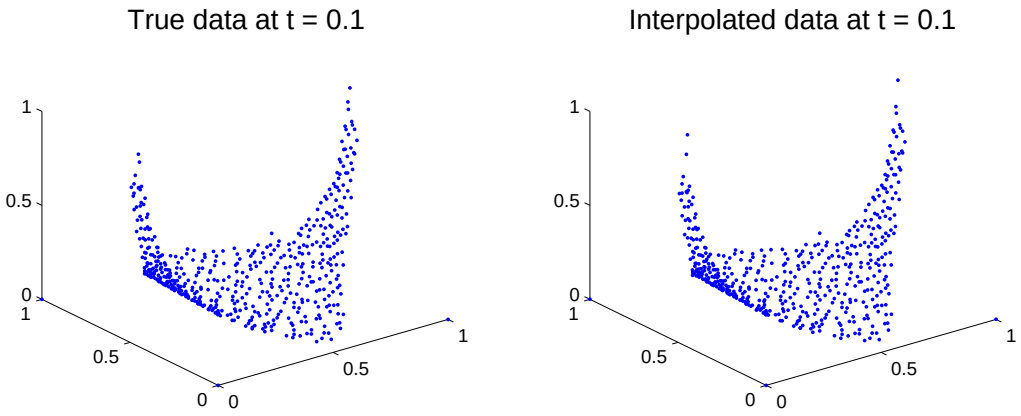

Figure 3.2: Left: true data; Right: interpolated data at time $t = 0.1$

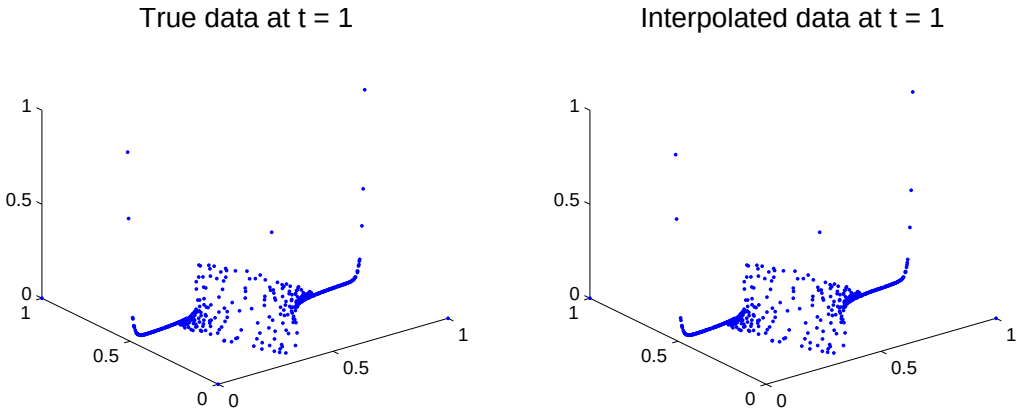

Figure 3.3: Left: true data; Right: interpolated data at time $t = 1.0$

Next, suppose we have the *Van der Pol* oscillator:

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -x_1 + (1 - x_1^2)x_2$$

$$\dot{\rho} = -\rho(1 - x_1^2).$$

Fig. 3.2 shows true and interpolated ρ values for 500 points based on $\hat{\boldsymbol{\alpha}}$ estimated

by the maximum-entropy basis functions of given other 500 points at time $t = 0.1$ and the Frobenius norm of the errors is 0.1925. And Fig. 3.3 shows the results at time $t = 1.0$. The result matches the true data well and the Frobenius norm of the errors is 0.0270.

4. NUMERICAL METHODS

In this chapter we briefly review the *method of characteristics* which is widely used to solve time-dependent partial differential equations. Then we present another new numerical method: the *Rothe & maximum-entropy method* which combines the maximum-entropy approximation with the Rothe method.

Given a deterministic dynamics $\boldsymbol{F}$ such that $\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x})$, the evolution of the probability density function $u_t = u(t, \cdot)$ in the presence of initial value uncertainties is described by the first-order linear partial differential equation [20]

$$\frac{\partial}{\partial t} u = \mathcal{A} u = -\mathrm{div}(\boldsymbol{F}\, u), \quad u(0, \cdot) = u_0 \tag{4.1}$$

where $\mathcal{A}$ is a differential operator involved only in spatial derivatives of u.

4.1 Method of Characteristics (MOC)

The method of characteristics first conducts a semi-discretization in space and as a result the problem is reduced to a system of ODEs. Solving these ODEs yields discrete solutions along trajectories in time.

A first-order PDE can be solved along characteristics which are the curves $(t(s), \boldsymbol{x}(s))_{s \in \mathbb{R}}$ describing the value $u(t(s), \boldsymbol{x}(s))$ of a solution u by an ODEs. The following shows that how the MOC transforms the Liouville equation to ODEs [20].

Since $\mathrm{div}(\boldsymbol{F}\, u) = \mathrm{div}(\boldsymbol{F})\, u + \langle \boldsymbol{F}, \nabla u \rangle$, (4.1) can be rewritten as

$$\frac{\partial u}{\partial t} + \sum_{i=1}^{n_t} F_i \cdot \frac{\partial u}{\partial x_i} = -\mathrm{div}(\boldsymbol{F})\, u \tag{4.2}$$

Letting $z(s) := u(t(s), \boldsymbol{x}(s))$ gives

$$\frac{dz}{ds} = \frac{d}{ds}u(t, \boldsymbol{x}) = \frac{\partial u}{\partial t} \cdot \frac{dt}{ds} + \sum_{i=1}^{n_t} \frac{\partial u}{\partial x_i} \cdot \frac{dx_i}{ds} \tag{4.3}$$

By setting $\frac{dt}{ds} = 1$ and $\frac{dx_i}{ds} = F_i$, we have

$$\frac{dz}{ds} = \frac{\partial u}{\partial t} + \sum_{i=1}^{n_t} F_i \cdot \frac{\partial u}{\partial x_i} = -\mathrm{div}(\boldsymbol{F}) \cdot z \tag{4.4}$$

Thus, the PDE (4.1) is transformed to the following ODEs

$$\begin{aligned} &\frac{d}{ds}t(s) = 1 \\ &\frac{d}{ds}\boldsymbol{x}(s) = \boldsymbol{F}(\boldsymbol{x}(s)) \\ &\frac{d}{ds}z(s) = -\mathrm{div}[\boldsymbol{F}(\boldsymbol{x}(s))] \cdot z(s) \end{aligned} \tag{4.5}$$

Given initial values $t(0) = 0,\ \boldsymbol{x}(0) = \boldsymbol{x}_0$ and $z(0) = u(0, \boldsymbol{x}_0)$, the solution of (4.1) is

$$u(t, \boldsymbol{x}(t)) = u(0, \boldsymbol{x}_0) \exp\left(- \int_0^t \mathrm{div}[\boldsymbol{F}(\boldsymbol{x}(s))] ds \right) \tag{4.6}$$

As described previously, solving the Liouville equation using the MOC yields discrete solutions along trajectories in time. For this reason the initial points given in a spatial domain will gather, in most cases of interest, or disperse over time depending on the given system and this makes it impossible to evaluate u_t over the specific spatial domain of our interest.

4.2 Rothe & Maximum-Entropy Method

The Rothe method first conducts a semi-discretization in time while the MOC does in space. By considering the time-dependent PDE as an ODE in a function

space, the problem is reduced to a stationary or time-independent PDE [7]. Solving this PDE yields an approximation of u_t at a discrete time.

4.2.1 Semi-discretization in Time

As defined in Chapter 2, the true solution of the probability density function at time t, u_t, is described by the semigroup of *Frobenius-Perron operators* $\{\mathcal{P}_t\}_{t\in[0,T]}$

$$u_t = \mathcal{P}_t u_0, \quad t \in [0, T] \tag{4.7}$$

and the evolution operator $\mathcal{P}_t$ is approximated using rational approximations $r(z)$ to e^z

$$\mathcal{P}_\tau \approx R_\tau := r(\tau\mathcal{A}) \tag{4.8}$$

For our work we approximate u_t using *A-stable rational approximations* $R_\tau^{(k)}$ to $\mathcal{P}_\tau$ [3]

$$\begin{aligned} u_{t_{j+1}} = R_\tau^{(k)} u_{t_j} &= R_\tau^{(k)} \left(R_\tau^{(k)} u_{t_{j-1}} \right) \\ &= \left(R_\tau^{(k)} R_\tau^{(k)} \cdots R_\tau^{(k)} \right) u_0 \end{aligned} \tag{4.9}$$

where $\tau > 0$ is a small time step such that $t_{j+1} = t_j + \tau$. This discrete evolution approximated at t_{j+1} by A-stable rational approximations has consistency order k which depends on the choice of a rational function. The definitions of A-stability and consistency order are as follows:

Definition 4.2.1 (A-stability): An approximation $r(z)$ to e^z is called *A-stable* if its stability region $\{z \in \mathbb{C},\, r(z) \leq 1\}$ includes all complex numbers with non-positive real part, i.e.,

$$|r(z)| \leq 1, \qquad \forall z \in \mathbb{C}, \quad \Re(z) \leq 0.$$

The semi-discretization scheme R_τ defined by $r(\tau\mathcal{A})$ is called A-stable if r is A-stable.

Definition 4.2.2 (Consistency order): The local error

$$\varepsilon_t(u,\tau) := \mathcal{P}_\tau u_t - R_\tau u_t$$

is called *consistency error* and the discrete evolution $R_\tau^j,\ j = 1,\dots,T/\tau$, is called *consistent* if for all $t \in [0,T]$

$$\lim_{\tau\to 0} ||\varepsilon_t(u,\tau)|| = 0.$$

The discrete evolution has *consistency order* k if for all $t \in [0,T]$

$$||\varepsilon_t(u,\tau)|| = \mathcal{O}(\tau^{k+1}) \quad \text{as} \quad \tau \to 0.$$

If the rational approximation $r(z)$ is consistent and A-stable, the convergence of the discrete evolution is guaranteed by the following theorem [3].

Theorem 4.2.3 (Rational approximation of semigroup): Let $\mathcal{A}$ generate a strongly continuous semigroup $\mathcal{P}_t = e^{t\mathcal{A}},\ t > 0$, i.e.,

$$\lim_{t\to 0} ||\mathcal{P}_t u - \mathcal{P}_0 u|| = 0, \quad \forall u \in D(\mathcal{P}_t)$$

and let further

$$||\mathcal{P}_t|| \le 1, \quad \forall t \ge 0 \tag{4.10}$$

Then for any A-stable rational approximation $r(z)$ to e^z of consistency order k there is a constant $c > 0$ such that for all $u \in D(\mathcal{A}^{k+1})$

$$||\mathcal{P}_{t_n} u - R_\tau^n u|| \le c\, t_n \tau^k ||\mathcal{A}^{k+1} u||, \quad t_n = n\,\tau,\ \tau > 0,\ n \in \mathbb{N}.$$

The semigroup of Frobenius-Perron operators $\mathcal{P}_t,\ t \geq 0$ is strongly continuous and satisfies (4.10) [14]. Thus, if the rational function satisfies *consistency order k* and *A-stability*, a discrete solution converges to the analytical solution with order k.

On choosing A-stable rational functions of consistency order k, the *Padé approximation* to the exponential function e^z defined by

$$r(z) = \frac{a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n}{b_0 + b_1 z + b_2 z^2 + \cdots + b_d z^d}$$

gives some choices.

$$r(z) = \frac{1}{1-z}, \quad r(z) = \frac{1+\frac{1}{2}z}{1-\frac{1}{2}z}, \quad r(z) = \frac{1+\frac{1}{3}z}{1-\frac{2}{3}z+\frac{1}{6}z^2}, \ \cdots$$

The first one is also known as the backward Euler method and yields consistency order 1, whereas the second one named the *Crank-Nicolson method* yields consistency order 2, which is used for our approach in this thesis. Indeed, the higher consistency order we use, the more accurate approximation we can achieve; however, the computation becomes very costly because of the second-order approximation $\mathcal{A}^2 u$. Thus, we choose the function of consistency order 2. The rational approximation defined in (4.9) with

$$R_\tau^{(k)} = r(\tau\mathcal{A}) = \frac{\mathrm{Id} + \frac{\tau}{2}\mathcal{A}}{\mathrm{Id} - \frac{\tau}{2}\mathcal{A}} \tag{4.11}$$

yields discrete solutions of consistency order 2 and we obtain the following time-independent or *stationary spatial problem*

$$\left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right) u_{t_{j+1}} = \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right) u_{t_j} \tag{4.12}$$

where Id is an identity operator.

Unlike the MOC, the Rothe method sticks to a fixed spatial domain over time so we can evaluate u_t over the specific spatial domain of our interest.

4.2.2 Solutions of the Stationary Spatial Problem

The stationary spatial problem previously proposed can be solved by the spatial discretization technique, that is, approximation. There are several approximation schemes such as radial basis functions, moving least-square approximants, natural neighbor-based interpolants, etc. In this thesis we use the entropy-based approximation scheme: *maximum-entropy approximation* for least-biased interpolations, which was introduced in the previous chapter.

For a fixed spatial domain, Ω, we have the same maximum-entropy basis functions $\phi_i(\boldsymbol{x})$ at any time step. Thus, $u_{t_{j+1}}$ and u_{t_j} are both approximated by $\phi_i(\boldsymbol{x})$

$$\hat{u}_{t_{j+1}}(\boldsymbol{x}) = \sum_{k=1}^{n} \alpha_{t_{j+1},k}\, \phi_k(\boldsymbol{x}) \tag{4.13}$$

$$\hat{u}_{t_j}(\boldsymbol{x}) = \sum_{k=1}^{n} \alpha_{t_j,k}\, \phi_k(\boldsymbol{x}) \tag{4.14}$$

where $\alpha_{t_{j+1},k}$, $\alpha_{t_j,k}$ are coefficients at $t = t_{j+1}$, $t = t_j$, respectively and they are to be determined later. Plugging (4.13) and (4.14) into (4.12) yields

$$\left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\left(\sum_{k=1}^{n} \alpha_{t_{j+1},k}\, \phi_k(\boldsymbol{x})\right) = \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\left(\sum_{k=1}^{n} \alpha_{t_j,k}\, \phi_k(\boldsymbol{x})\right) \tag{4.15}$$

Since $\mathcal{A}$ is linear and the sum over $\boldsymbol{x}_i$ is finite, this equation is equivalent to

$$\sum_{k=1}^{n} \alpha_{t_{j+1},k}\left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_k(\boldsymbol{x}) = \sum_{k=1}^{n} \alpha_{t_j,k}\left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_k(\boldsymbol{x}) \tag{4.16}$$

With the maximum-entropy basis functions, we can analytically express the action

of the differential operator $\mathcal{A}$ of the semigroup of Frobenius-Perron operators on the basis functions by the following equation

$$\mathcal{A}\phi = -\mathrm{div}(\boldsymbol{F})\,\phi - \langle \boldsymbol{F}, \nabla\phi \rangle \tag{4.17}$$

By evaluating (4.16) over all the given points $\boldsymbol{x}_i$ using (4.17), it can be restated as a system of linear equations

$$\boldsymbol{A}\,\boldsymbol{\alpha}_{t_{j+1}} = \boldsymbol{B}\,\boldsymbol{\alpha}_{t_j} \tag{4.18}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ are defined as

$$\boldsymbol{A} = \begin{bmatrix} \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_1(\boldsymbol{x}_1) & \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_2(\boldsymbol{x}_1) & \dots & \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_n(\boldsymbol{x}_1) \\ \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_1(\boldsymbol{x}_2) & \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_2(\boldsymbol{x}_2) & \dots & \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_n(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_1(\boldsymbol{x}_n) & \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_2(\boldsymbol{x}_n) & \dots & \left(\mathrm{Id} - \frac{\tau}{2}\mathcal{A}\right)\phi_n(\boldsymbol{x}_n) \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_1(\boldsymbol{x}_1) & \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_2(\boldsymbol{x}_1) & \dots & \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_n(\boldsymbol{x}_1) \\ \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_1(\boldsymbol{x}_2) & \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_2(\boldsymbol{x}_2) & \dots & \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_n(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_1(\boldsymbol{x}_n) & \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_2(\boldsymbol{x}_n) & \dots & \left(\mathrm{Id} + \frac{\tau}{2}\mathcal{A}\right)\phi_n(\boldsymbol{x}_n) \end{bmatrix}$$

The coefficients $\boldsymbol{\alpha}_{t_{j+1}}$ and $\boldsymbol{\alpha}_{t_j}$ are the unknowns to be determined in the following

where

$$\boldsymbol{\alpha}_{t_{j+1}} = [\alpha_{t_{j+1},1} \;\; \alpha_{t_{j+1},2} \;\; \cdots \;\; \alpha_{t_{j+1},n}]^{\mathrm{T}}$$
$$\boldsymbol{\alpha}_{t_j} = [\alpha_{t_j,1} \;\; \alpha_{t_j,2} \;\; \cdots \;\; \alpha_{t_j,n}]^{\mathrm{T}}$$

Since u_{t_j} is given at each time step, t_{j+1}, we can solve for $\boldsymbol{\alpha}_{t_j}$ from (4.14):

$$\boldsymbol{\alpha}_{t_j} = \boldsymbol{\Phi}^{-1}\, \boldsymbol{u}_{t_j} \tag{4.19}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_1(\boldsymbol{x}_1) & \phi_2(\boldsymbol{x}_1) & \dots & \phi_n(\boldsymbol{x}_1) \\ \phi_1(\boldsymbol{x}_2) & \phi_2(\boldsymbol{x}_2) & \dots & \phi_n(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{x}_n) & \phi_2(\boldsymbol{x}_n) & \dots & \phi_n(\boldsymbol{x}_n) \end{bmatrix}, \quad \boldsymbol{u}_{t_j} = \begin{bmatrix} u_{t_j}(\boldsymbol{x}_1) \\ u_{t_j}(\boldsymbol{x}_2) \\ \vdots \\ u_{t_j}(\boldsymbol{x}_n) \end{bmatrix}$$

Therefore, combining (4.19) with (4.18) yields the solution for $\boldsymbol{\alpha}_{t_{j+1}}$

$$\boldsymbol{\alpha}_{t_{j+1}} = \boldsymbol{A}^{-1}\, \boldsymbol{B}\, \boldsymbol{\Phi}^{-1}\, \boldsymbol{u}_{t_j} \tag{4.20}$$

and this enables us to compute $\boldsymbol{u}_{t_{j+1}}$

$$\boldsymbol{u}_{t_{j+1}} = \boldsymbol{\Phi}\, \boldsymbol{\alpha}_{t_{j+1}} \tag{4.21}$$

One of benefits with this method is that for a fixed spatial domain, the matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{\Phi}$ do not change at each time step. Once these matrices are computed at the beginning, no more computations are required. Thus, it vastly reduces the total computation time.

4.3 Numerical Examples

Consider a two-state Markov jump linear system which is defined by

$$\boldsymbol{x}_{k+1} = A_\sigma \boldsymbol{x}_k, \quad \sigma \in \{1, 2\}$$
$$A_1 = \begin{bmatrix} 0.7 & 0 \\ 0 & 1 \end{bmatrix}, \; A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0.85 \end{bmatrix}$$

and whose initial condition is given by

$$\boldsymbol{x}_0 \sim \mathcal{N}(\mu_0, \Sigma); \quad \mu_0 = [7, 7]^T, \; \Sigma = \begin{bmatrix} 1.5^2 & 0 \\ 0 & 1.5^2 \end{bmatrix}$$

Here the Markov jump linear system is defined as follows:

Definition 4.3.1 (Markov Jump Linear System): The discrete-time Markov jump linear systems is defined by [5]

$$\boldsymbol{x}(k+1) = A_{\sigma_k} \boldsymbol{x}(k), \quad k \in \mathbb{N}_0 \tag{4.22}$$

where the state $\boldsymbol{x}(k) \in \mathbb{R}^n$, the system matricx $A_{\sigma_k} \in \mathbb{R}^{n \times n}$ and $\sigma_k \in \mathcal{M} := \{1, 2, \ldots, m\}$. σ_k is the discrete-time Markov chain with a mode transition probability

$$\mathbb{P}(\sigma_{k+1} = j \mid \sigma_k = i) = p_{ij} \tag{4.23}$$

where $p_{ij} \geq 0, \; \forall i, j \in \mathcal{M}$. The probability distribution $\pi(k) \in \mathbb{R}^m$ of (4.22) is governed by

$$\pi(k+1) = \pi(k)P, \quad \pi(0) = [\pi_1(0), \cdots, \pi_m(0)] \tag{4.24}$$

where $P \in \mathbb{R}^{m \times m}$ denotes the transition probability matrix and its row sum $\sum_{j=1}^{m} p_{ij} =$

$1, \forall i \in \mathcal{M}$.

Fig. 4.1 shows that how the probability density function of this system with the initial condition given by the Gaussian evolves through the Rothe & maximum-entropy method under the Markov switching with an initial probability π_0 and transition probability matrix P given by

$$\pi_0 = [0.5, 0.5], \quad P = \begin{bmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{bmatrix}$$

In the simulation 3,000 scattered points in $[0, 14]^2$ are used and the Markov jump is set to occur every 1 second; between each jump the system is governed by each system matrix A_1 and A_2. As a comparison, Fig. 4.2 shows how the probability density function of the same system evolves through the MOC using 10,000 points in $[0, 14]^2$.

For tracking the PDF propagation of the switched system, it is required to know

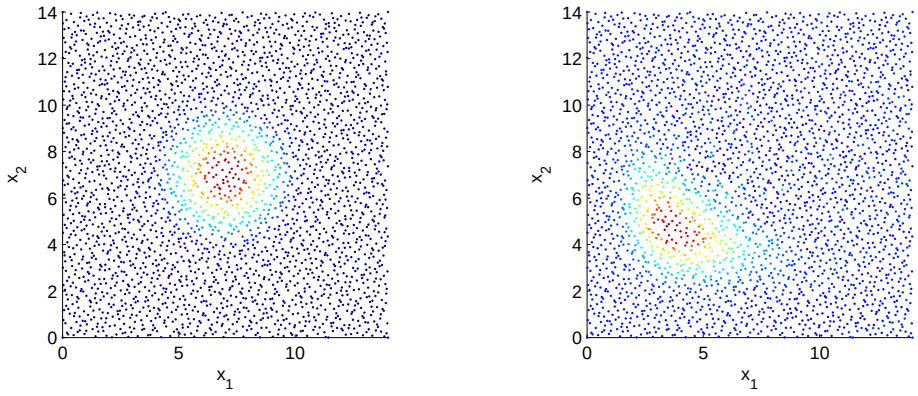

Figure 4.1: Rothe & Maximum-Entropy Method: PDF propagation of 3,000 scattered points for Markov Jump Linear System (Left: t = 0, Right: t = 5)

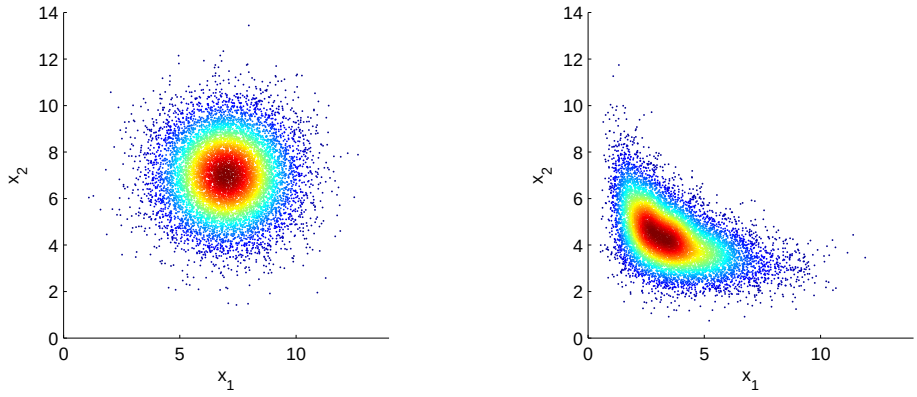

Figure 4.2: Method of Characteristics: PDF propagation of 10,000 points for Markov Jump Linear System (Left: t = 0, Right: t = 5)

PDFs of all the modes, two modes in this example, at every switching; however, for the MOC the PDF information is only available for the switched-on mode each switching time since the points move along the trajectories as depicted in Fig. 4.2. For this reason the MOC needs to interpolate PDFs of the switched-off modes based on PDFs of the switched-on mode at every switching, which causes inaccuracy in tracking the PDF propagation. Whereas, the Rothe & maximum-entropy method plays in the fixed spatial domain over time as depicted in Fig. 4.1 and this enables us to capture PDFs of all the modes at every switching without interpolations. As a result, more accuracy is guaranteed.

For the next example, consider the following Michaelis-Menten kinetics:

$$\dot{x}_1 = -\frac{x_2}{K_m + x_1}\, x_1$$

$$\dot{x}_2 = 0$$

where $K_m = 1$ and whose initial condition is given by

$$\boldsymbol{x}_0 \sim \mathcal{N}(\mu_0, \Sigma); \quad \mu_0 = [2, 2]^T, \; \Sigma = \begin{bmatrix} 1/8 & 0 \\ 0 & 1/40 \end{bmatrix}$$

Fig. 4.3 shows the evolution of the probability density function of this system through the Rothe & maximum-entropy method. Also, it is shown that the total probability mass is preserved over time.

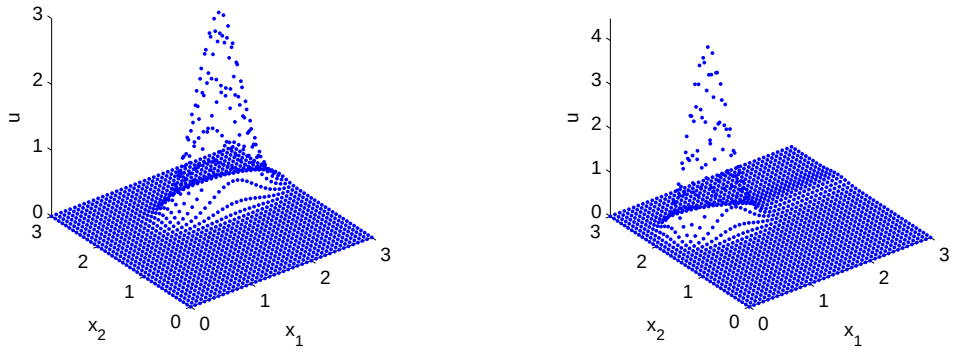

Figure 4.3: Rothe & Maximum-Entropy Method: PDF propagation of 1,600 points for the Michaelis-Menten kinetics (Left: t = 0, Right: t = 1)

5. CONCLUSION

In this thesis we developed a new method, the Rothe & maximum-entropy method, for analyzing dynamical systems with initial value uncertainties. To this end, we first conducted a semi-discretization in time using the Rothe method and this yielded a time-independent or stationary spatial problem. Then we solved the stationary spatial problem using the spatial discretization technique, the maximum-entropy approximation. Through these steps we finally obtained a system of linear equations for the probability density function at time t, u_t.

In the numerical examples it is shown that this method is capable of evaluating the evolution of probability density functions with even fewer number of scattered points over the specific fixed domain of interest without moving the domain. Especially for the switched system, this method provided both modes with seamless information on PDFs over the same spatial substrate regardless of switching. As a result, we were able to accurately capture the PDF propagation without interpolating PDFs of the switched-off mode based on PDFs of the switched-on mode, which is used for the method of characteristics and may result in inaccurate PDF propagation.

Compared to the existing methods: the method of characteristics which is widely used and the classical Rothe method which uses grid points, the Rothe & maximum-entropy method offers the following advantages:

1. Computational cost is low compared to the classical Rothe method using grid points, especially for higher dimensional systems since the spatial discretization is performed by the maximum-entropy approximation which enables unbiased interpolations even with fewer number of scattered points.

2. Monitoring the evolution of the probability density functions in the fixed do-

main of our interest, that is, at the same points over time is possible since the Rothe method enables us to keep the spatial domain without changing it.

3. With these two advantages above, we may be able to find invariant sets for dynamical systems with initial value uncertainties.

REFERENCES

[1] N. Agmon, Y. Alhassid, and R.D. Levine. An algorithm for finding the distribution of maximal entropy. *Journal of Computational Physics*, 30(2):250–258, 1979.

[2] Y. Alhassid, N. Agmon, and R.D. Levine. An upper bound for the entropy and its applications to the maximal entropy problem. *Chemical Physics Letters*, 53(1):22–26, 1978.

[3] P. Brenner and V. Thomée. On rational approximations of semigroups. *SIAM Journal on Numerical Analysis*, 16(4):683–694, 1979.

[4] P. Deuflhard and F. Bornemann. *Scientific computing with ordinary differential equations*, volume 42. New York: Springer-Verlag, 2002.

[5] O.L. do Valle Costa, M.M.D. Fragoso, and R.P. Marques. *Discrete time Markov jump linear systems*. London: Springer-Verlag, 2005.

[6] P. Dutta, A. Halder, and R. Bhattacharya. Uncertainty quantification for stochastic nonlinear systems using perron-frobenius operator and karhunen-loève expansion. In *Control Applications (CCA), 2012 IEEE International Conference on*, pages 1449–1454. IEEE, 2012.

[7] C. Grossmann and H. Roos. *Numerical treatment of partial differential equations*. Berlin: Springer-Verlag, 2007.

[8] A. Halder and R. Bhattacharya. Dispersion analysis in hypersonic flight during planetary entry using stochastic liouville equation. *Journal of Guidance, Control, and Dynamics*, 34(2):459–474, 2011.

[9] A. Halder, K. Lee, and R. Bhattacharya. Probabilistic robustness analysis of f-16 controller performance: An optimal transport approach. *Parameters*, 2:1d.

[10] E.T. Jaynes. Information theory and statistical mechanics. *Physical review*, 106(4):620, 1957.

[11] E.T. Jaynes. Information theory and statistical mechanics. ii. *Physical review*, 108(2):171, 1957.

[12] O. Junge, J.E. Marsden, and I. Mezic. Uncertainty in the dynamics of conservative maps. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, volume 2, pages 2225–2230. IEEE, 2004.

[13] A. Lasota. *Probabilistic properties of deterministic systems*. Cambridge, UK: Cambridge University Press, 1985.

[14] A. Lasota and M.C. Mackey. *Chaos, fractals, and noise: stochastic aspects of dynamics*, volume 97. New York: Springer-Verlag, 1994.

[15] C.E. Shannon, W. Weaver, R.E. Blahut, and B. Hajek. *The mathematical theory of communication*, volume 117. Urbana, IL: University of Illinois Press, 1949.

[16] G. Strang and G.J. Fix. *An analysis of the finite element method*, volume 212. Englewood Cliffs, NJ: Prentice-Hall, 1973.

[17] N. Sukumar. Construction of polygonal interpolants: a maximum entropy approach. *International Journal for Numerical Methods in Engineering*, 61(12):2159–2181, 2004.

[18] N. Sukumar and E.A. Malsch. Recent advances in the construction of polygonal finite element interpolants. *Archives of Computational Methods in Engineering*, 13(1):129–163, 2006.

[19] N. Sukumar and R.W. Wright. Overview and construction of meshfree basis functions: from moving least squares to entropy approximants. *International Journal for Numerical Methods in Engineering*, 70(2):181–205, 2007.

[20] A.Y. Weiße. *Global Sensitivity Analysis of Ordinary Differential Equations*. PhD thesis, FU Berlin, 2009.