# IDENTIFICATION OF INPUT-FREE FINITE LATTICE DYNAMICAL SYSTEMS UNDER ENVELOPE CONSTRAINTS

J. Barrera, P. Martin, M. Gubitoso, N. Hirata*, N. Trepode†
*Institute of Mathematics and Statistics*
*University of São Paulo*
*Rua do Matão, 1010 – CEP 05508-090 – São Paulo, Brazil*

E. Dougherty
*Department of Electrical Engineering*
*Texas A&M University*
*College Station - 77843-3128 - TX - United States*

**Abstract** In this paper, we call finite dynamical system (FDS) a discrete time and finite range dynamical system. A FDS is called a finite lattice dynamical system (FLDS) when its transition function is defined on a finite lattice. Hence, the transition function of a FLDS is a lattice operator and can be represented by an union of sup-generating operators, characterized by the operator basis. The identification of a FLDS consists in the determination of the transition function basis from samples of the system dynamics. In this paper, we introduce the notion of dynamical envelope constraint (i.e., a lower and upper limit for the dynamics of the system to be identified) and show how to apply it to improve the precision of the identification.

**Keywords:** finite lattice dynamical systems, system identification, envelope constraint.

## 1. Introduction

In the automatic control literature, we find several studies on identification of discrete time dynamical systems (usually linear). In this paper, we are

*Partially supported by CNPq - Brazil, proc. 68.0093/01-0 and 300238/01-0.
†Supported by grant 00/10684 from FAPESP - Brazil.


Redistribution rights reserved CSIRO Publishing. ISBN 0 643 06804 X

concerned with the identification of discrete time and finite range dynamical (non linear) systems.

We represent *discrete time* by an integer number. A *signal* is a function from the set of integers to a finite set. A *Finite Dynamical System (FDS)* is a set of ordered pairs of signals, called *system dynamics*. The left element of the pair is the *state trajectory*, while the right element is the corresponding *input*. A system dynamics shows how the system states change in time under a given input. If a FDS is invariant under time translation, then it is called *translation invariant*.

An *input free FDS* is a set of signals representing the state trajectories, that is, state trajectory and dynamics are equivalent in input free systems. Translation invariant input free systems are uniquely characterized by a mapping in the space of states: the *transition function*. The system dynamics is generated by iterations of the transition function. In input free Finite Lattice Dynamical Systems (FLDS) the transition function is a lattice operator [5, 2].

Learning a concept (i.e., a function $\phi$) in a space of hypothesis (i.e., set of candidate functions represented in a given algebraic structure) consists in choosing a hypothesis that best agrees with a set of examples (i.e., pairs $(x_i, \phi(x_i))$), extracted from a random domain of unknown distribution [1, 4]. The error of the concept learned (i.e., a statistical distance between the target concept and the hypothesis learned) depends on the amount of available examples. For a fixed hypothesis space and a fixed domain distribution, the error decreases when the amount of available examples increases. However, usually the amount of examples available is not enough to get small errors. One way to approach this problem is by using prior information for reducing the hypothesis space. In fact, for a fixed domain distribution, if the hypothesis space is a subset of the original hypothesis space that includes the target concept (or, at least, a concept similar to the target concept), then the amount of examples necessary to get low errors is smaller than the one needed in the original space.

The learning of a transition function from a random sample of the system dynamics is called *system identification*. In this learning process there are some difficulties: (*i*) cumulative errors due to the iteration of the learned transition function; (*ii*) constraining the hypothesis space (i.e., the set of candidate transition functions) depends on prior knowledge about the target system dynamics and on transition functions families that generate dynamics consistent with that.

In this paper, we study the identification of input-free FLDS under dynamical envelope constraints, that is, the hypothesis space considered is such that the dynamics generated by iteration of a hypothesis is always between the dynamics generated by two limiting hypothesis: the lower and upper dynamical extremities.

Following this introduction, Section 2 recalls the concept of FDS and the representation of FLDS. Section 3 presents a model for identification of FDS. Section 4 presents the technique of FLDS identification under envelope constraints. Section 5 gives experimental results, comparing envelope constrained and unconstrained system identification. Section 6 discusses the results obtained and proposes some future steps for this research.

## 2. Finite lattice dynamical system

Let $X$ be a finite set, $k$ be a positive integer, and $\phi$ be a mapping from $X^k$ to $X$. *The input free Finite Dynamical System (FDS)* $S(\phi)$ is given by, for every $i \geq k$,

$$x_{i+1} = \phi(x_i, x_{i-1}, \ldots, x_{i-k})$$

where $x_j \in X$, for all $j \geq 0$. $\phi$ is called the *transition function.* The variables $i$, $x$ and $k$ are called, respectively, *discrete time*, *state* and *system memory.* The set $X$ is called the *state space.*

Systems defined as above are time translation invariant, that is, the transition function is the same for all discrete time $i \geq k$.

*A FDS* $S(\phi)$ can be represented equivalently by

$$x_{i+1} = \varphi(x_i),$$

where $x_j \in X^{k+1}$ and $\varphi$ is a mapping defined on $X^{k+1}$.

In other words, a FDS $S(\phi)$ of memory $k$ in the space $X$ can be represented by a FDS $S(\varphi)$ of memory 1 in the space $X^{k+1}$. Hence, without loss of generality, we can study just FDS of memory 1.

The state $x_0$ is called an *initial condition* and the sequence $x_0, \varphi(x_0), ..., \varphi^n(x_0)$ is a *dynamics* of the FDS $S(\varphi)$.

If the state space $X$ is a finite lattice, then the transition function $\phi$ is a lattice operator and can be represented by a union of sup-generating operators, characterized by the operator basis [3]. In this case, $S(\phi)$ is called an *input free Finite Lattice Dynamical System* (FLDS). In particular, if $X$ is given by $(n-1)$ Cartesian products of a finite chain (i.e., a finite completely ordered set) $K$, that is, $X = K^n$, then the transition function $\phi$ can be decomposed in a vector of $n$ transition functions, with components $\phi_j$ from $X$ to $K$. Once $\phi_j$ is also a lattice operator, it has a basis and a corresponding representation in terms of sup-generating operators. Therefore, the representation of the transition function $\phi$ is parameterized by the set of basis $\{\mathbf{B}(\phi_j) : j \in \{1, 2, ..., n\}\}$.

## 3. System Identification

The identification of a FDS $S(\psi)$ consists in applying an algorithm $L$, the *identification algorithm*, that constructs a good approximation $S(\phi)$ for $S(\psi)$ from samples of the system dynamics. For practical purposes, $L$ constructs the transition function $\phi$ in a convenient algebraic representation. When $S(\phi)$ is a FLDS, usually $L$ gives the basis of the transition function $\phi$.

Suppose that the state space $X$ is structured according with some fixed, but unknown, distribution $\mu$. A *training sample* $s$ for a *target* FDS $S(\psi)$ is generated by drawing from $X$, according with $\mu$, a set of random independent initial conditions $x_1, \ldots, x_m$. For each such initial condition $x_j$, we sample the next $k$ iterations of $\psi$, that is,

$$s(\psi, x_j) = \{x_j, \psi(x_j), \psi^{(2)}(x_j), \ldots, \psi^{(k)}(x_j)\}.$$

Let $l$ denote a *loss function*, that is, a function from $X^k \times X^k$ to $[0, +\infty)$ that measures the distance between dynamics in $X^k$. The *simple cost* loss function

$l_{SC}$ is the one that counts the number of corresponding differences in the two dynamics compared, that is, for every $u, v \in X^k$,

$$l_{SC}(u,v) = |\{u_j : u_j \neq v_j, j \in \{1,..,k\}\}|,$$

where $u_j$ and $v_j$ denote, respectively, the components of $u$ and $v$.

We assume that the transition function of the target FDS $\psi$ belongs to some hypothesis space $H \subseteq X^X$. Given the transition function of a target system $\psi \in H$, the *error* of any hypothesis $\phi \in H$ will be

$$Er_\mu(\phi,\psi,k) = \sum_{x\in X} l(s(\psi,x),s(\phi,x))\, P(s(\psi,x),s(\phi,x)),$$

where $P(s(\psi,x),s(\phi,x))$ is the join probability of $s(\psi,x)$ and $s(\phi,x)$. In the present formulation, with a deterministic target dynamical system, $P(s(\psi,x),s(\phi,x)) = \mu\{x\}$.

The error for the simple cost loss function is given, equivalently, by

$$Er_\mu(\phi,\psi,k) = \sum_{j=1}^{k} \mu\{x \in X \ : \ \phi^{(j)}(x) \neq \psi^{(j)}(x)\},$$

where $k$ can vary from 1 to $|X|$.

For simplicity, if $k = |X|$, we will denote just $Er_\mu(\phi,\psi)$. It is clear that if $k_1 \leq k_2$ then $Er_\mu(\phi,\psi,k_1) \leq Er_\mu(\phi,\psi,k_2)$, and so increasing $k$ permits a kind of refinement of the error measure.

**Example.** Let the state space be $X = \{1,2,3,4,5\}$, the transition function of the target system be $\psi = (2,3,1,5,4)$, that is, $\psi(1) = 2$, $\psi(2) = 3$, ..., $\psi(5) = 4$, and suppose that $\mu$ is a uniform distribution on $X$. If the transition function of the hypothesis system is $\phi = (4,3,1,5,2)$, then we have:

| | | |
|---|---|---|
| $\psi = (2,3,1,5,4)$ | $\phi = (4,3,1,5,2)$ | $\mu\{x \ : \ \psi(x) \neq \phi(x)\} = 2/5$ |
| $\psi^2 = (3,1,2,4,5)$ | $\phi^2 = (5,1,4,2,3)$ | $\mu\{x \ : \ \psi^2(x) \neq \phi^2(x)\} = 4/5$ |
| $\psi^3 = (1,2,3,5,4)$ | $\phi^3 = (2,4,5,3,1)$ | $\mu\{x \ : \ \psi^3(x) \neq \phi^3(x)\} = 1$ |
| $\psi^4 = (2,3,1,4,5)$ | $\phi^4 = (3,5,2,1,4)$ | $\mu\{x \ : \ \psi^4(x) \neq \phi^4(x)\} = 1$ |
| $\psi^5 = (3,1,2,5,4)$ | $\phi^5 = (1,2,3,4,5)$ | $\mu\{x \ : \ \psi^5(x) \neq \phi^5(x)\} = 1$ |

and so the error is $Er_\mu(\phi,\psi) = \frac{21}{5}$. If the transition function of the hypothesis is $\varphi = (3,1,2,5,4)$, we have

| | | |
|---|---|---|
| $\psi = (2,3,1,5,4)$ | $\varphi = (3,1,2,5,4)$ | $\mu\{x \ : \ \psi(x) \neq \varphi(x)\} = 3/5$ |
| $\psi^2 = (3,1,2,4,5)$ | $\varphi^2 = (2,3,1,4,5)$ | $\mu\{x \ : \ \psi^2(x) \neq \varphi^2(x)\} = 3/5$ |
| $\psi^3 = (1,2,3,5,4)$ | $\varphi^3 = (1,2,3,5,4)$ | $\mu\{x \ : \ \psi^3(x) \neq \varphi^3(x)\} = 0$ |
| $\psi^4 = (2,3,1,4,5)$ | $\varphi^4 = (3,1,2,4,5)$ | $\mu\{x \ : \ \psi^4(x) \neq \varphi^4(x)\} = 3/5$ |
| $\psi^5 = (3,1,2,5,4)$ | $\varphi^5 = (2,3,1,5,4)$ | $\mu\{x \ : \ \psi^5(x) \neq \varphi^5(x)\} = 3/5$ |

and so the error is $Er_\mu(\varphi,\psi) = \frac{12}{5}$, which is almost half of the above error.

The tables above show that the hypothesis $\phi$ is "nearer" to $\psi$ than the hypothesis $\varphi$, but the system generated by $\phi$ is almost twice worse than the system generated by $\varphi$. It is interesting to note that

$$Er_{\mu=}(\phi,\psi,2) = Er_\mu(\varphi,\psi,2) = \frac{6}{5},$$

and so, up to the second iteration the error function does not differentiate $\phi$ and $\varphi$. $\circ$

Let $S(mk, \psi)$ denote the set of training samples of $m$ initial conditions with iteration blocks of length $k$, for a given FDS with transition function $\psi$. An element $s \in S(mk, \psi)$ is of the form

$$s = ((x_1, \psi(x_1), \ldots, \psi^{k-1}(x_1)), \ldots, (x_m, \psi(x_m), \ldots, \psi^{k-1}(x_m))).$$

It is clear that we have a natural bijection

$$\theta \colon X^m \longrightarrow S(mk, \psi)$$

given by $(x_1, \ldots, x_m) \mapsto s$. We will simply measure the set of samples $s \in S(mk, \psi)$ with property $P$ by $\mu^m(\theta^{-1}\{s \in S(mk, \psi) \ : \ s \text{ has property } P\})$.

An *identification algorithm* $L$ from $S(mk, \psi)$ to $H$ is called *Probably Approximately Correct (PAC)* if, for a fixed $k \geq 2$, there exists $m \geq m_0(\epsilon, \delta)$ such that

$$\mu^m(\theta^{-1}\{s \in S(mk, \psi) : \frac{1}{k} Er_\mu(L(s), \psi, k) < \epsilon \}) > 1 - \delta,$$

for any real numbers $\epsilon, \delta \in (0, 1)$, distribution $\mu$ on $X$, and target FDS $S(\psi)$, with $\psi \in H$. Note that term $\frac{1}{k}$ normalizes the error, that is, $0 \leq \frac{1}{k} Er_\mu(L(s), \psi, k) \leq 1$.

An identification algorithm $L$ for the hypothesis space $H$ is consistent if, given any training sample $s$ for a target FDS $S(\psi)$, with $\psi \in H$, the output hypothesis FDS $S(\phi)$ is such that $\phi = L(s) \in H$ agrees with $\psi$ on $s$.

If $H$ is a finite hypothesis space and $L$ is a consistent identification algorithm, then $L$ is PAC and

$$m_0(\epsilon, \delta) = \frac{(k-1)}{\epsilon} \ln(\frac{|H|}{\delta}).$$

Note that for $k = 2$, the last identification complexity formula reduces to the classical complexity expression of PAC learning. This formula is far from being a refined bound, due to the quite general conditions in which it was derived, but it shows clearly that the identification result becomes poorer when the hypothesis space increases. Thus, one way of simplifying the identification problem is choosing a smaller hypothesis space that includes the target FDS or, at least, a FDS very similar to the target. In the next section, we present a technique for constraining the hypothesis space: the dynamical envelopes.

## 4. Identification under envelope constraint

Let $X$ be a finite complete lattice, with partial order relation $\leq$, infimum $\wedge$ and supremum $\vee$. Yet, let $\alpha$ and $\beta$ be two operators defined on $X$. The pair of lattice operators $(\alpha, \beta)$ constitutes a *dynamical envelope* if and only if (iff), for every $i \geq 1$ and $x \in X$, $\alpha^i(x) \leq \beta^i(x)$.

A lattice operator $\phi$ is called *increasing* iff, for every $x, y \in X$, $x \leq y$ implies that $\phi(x) \leq \phi(y)$. The next proposition gives two sufficient conditions to build dynamical envelopes.

**Proposition 1** *Let $\alpha$ and $\beta$ be two lattice operators. The pair $(\alpha,\ \beta)$ constitutes a dynamical envelope if one of the following conditions hold:*

*1 $\alpha \leq \beta$ and $\beta$ is increasing;*

*2 $\alpha \leq \beta$ and $\alpha$ is increasing.*

***Proof 2*** *Let us prove 1. As $\alpha \leq \beta$, we have that $\alpha(x_0) \leq \beta(x_0)$ and, since $\beta$ is increasing, $\beta(\alpha(x_0)) \leq \beta(\beta(x_0))$. Again, since $\alpha \leq \beta$, we have that $\alpha(\alpha(x_0)) \leq \beta(\alpha(x_0)) \leq \beta(\beta(x_0))$. Hence, $\alpha^2(x_0) \leq \beta^2(x_0)$. This is the initial step of the proof by induction. The induction step (i.e., $\alpha^{i-1}(x_0) \leq \beta^{i-1}(x_0)$ implies that $\alpha^i(x_0) \leq \beta^i(x_0)$) is proved by the same arguments. By similar arguments, 2 can be proved.* ∘

Let $\alpha$ and $\beta$ be elements of the hypothesis space $H$ such that $(\alpha, \beta)$ is a dynamical envelope. $(\alpha, \beta)$ defines a subset $H'$ of the hypothesis space $H$ such that $\phi \in H' \Leftrightarrow \alpha^i \leq \phi^i \leq \beta^i$, for every $i \geq 1$.

The identification of a target FDS $S(\psi)$, with $\psi \in H$, constrained to $H'$, can be implemented as a two steps process: *i* - estimate the hypothesis $\phi \in H$ for $\psi$; *ii* - build the hypothesis $\theta$ in $H'$ through the projection of $\phi$ in $H'$, that is, for every $i \geq 1$,

$$\theta^i = (\phi^i \wedge \beta^i) \vee \alpha^i.$$

## 5. Experimental results

Let $K = [-100, ..., 0, ..., 100]$ be the range and $X = K^2$ be the state space. The lattice operators $\psi, \alpha$ and $\beta$ defined from $X^2$ to $X$ are the transition functions of, respectively, the target FDS, the lower and the upper extremities of the envelope. Figure 1 shows a simulation of these three systems. The two high frequency curves are $\psi_1$ and $\psi_2$, the components of $\psi$. The lower and the upper curves are, respectively, $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$, the components of the lower and upper extremities of the dynamical envelope.

We have chosen 15 initial conditions and simulated the target FLDS by 120 steps for each initial condition. After that, we extracted 50 examples of the type $(x_i, \psi(x_i))$ and learned $\phi$ by using the gray-scale Incremental Splitting of Intervals algorithm (GISI) [6], that generates a minimal collection of maximal intervals representing the basis of the components of $\phi$. Then, the dynamics $\phi^i$ and $(\phi^i \wedge \beta^i) \vee \alpha^i$ were simulated.

Finally, considering the *mean square error* (MSE) loss function, the error between $\psi^i$ and $\phi^i$, and between $\psi^i$ and $(\phi^i \wedge \beta^i) \vee \alpha^i$ is estimated. The simulation of all the 15 initial conditions are used for that.

Figure 2 shows the error curve for the two hypothesis, $\phi^i$ and $(\phi^i \wedge \beta^i) \vee \alpha^i$, estimated with samples of increasing size. Note that for small samples the error of $(\phi^i \wedge \beta^i) \vee \alpha^i$ is smaller than the error of $\phi^i$. However, for large samples, the situation is inverted.

This experiment was repeated for the same target FLDS $\psi$, but with a more severe dynamical envelope. Figures 3 and 4 give, respectively, the simulation

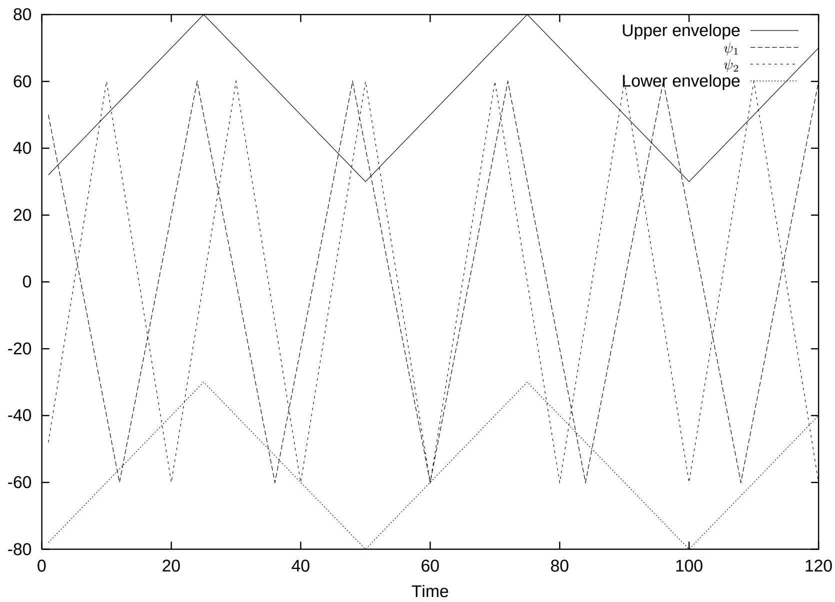


*Figure 1.* System 1 simulation.

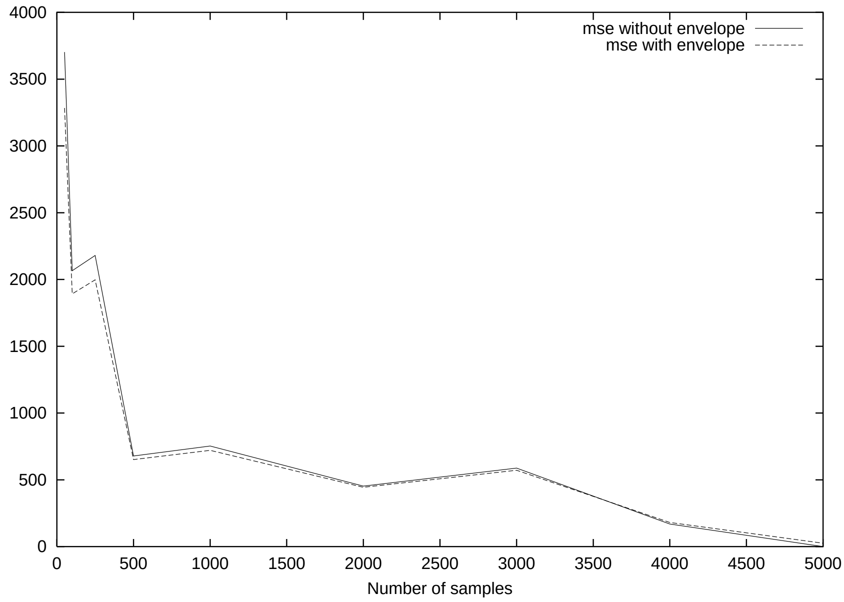


*Figure 2.* Error curves for system 1.

of these systems and the corresponding error curves. Observe that the same kind of phenomena observed in the first experiment happens again, but this time the error differences between $\phi^i$ and $(\phi^i \wedge \beta^i) \vee \alpha^i$ are more significant.

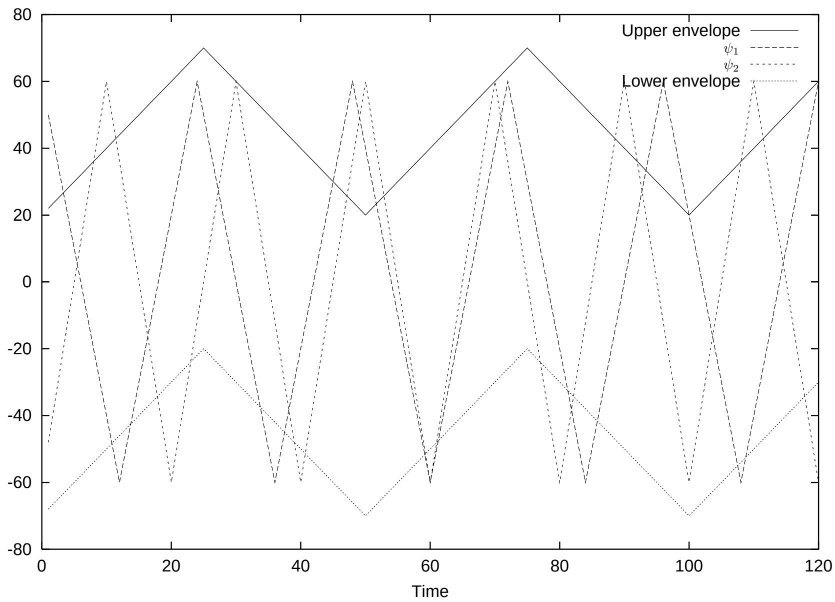


*Figure 3.* System 2 simulation.

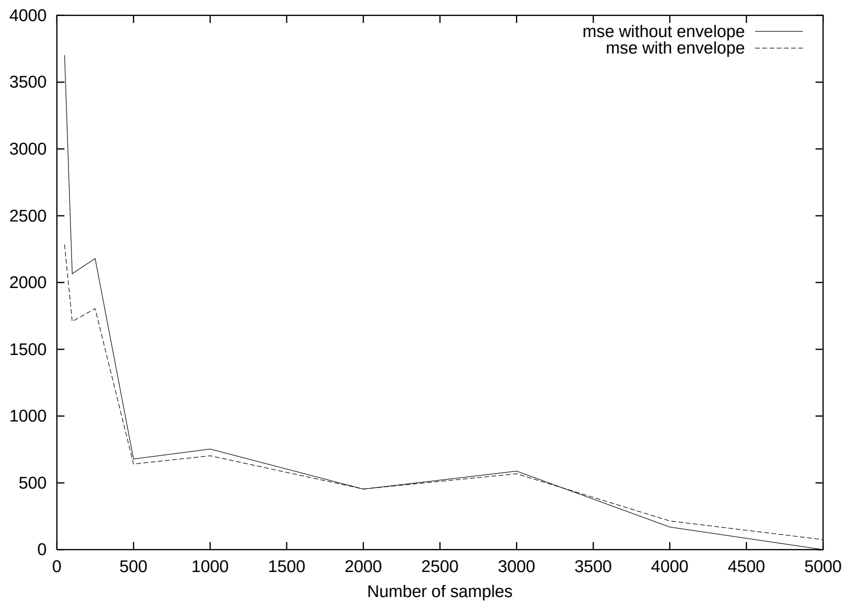


*Figure 4.* Error curves for system 2.

## 6. Conclusion

We presented the input free FLDS and studied their identification under a dynamical envelope constraint. Experimental results showed that envelope constraint may be beneficial for small samples.

In this system identification technique, the lattice structure is important for: *i* - representing the components of the transition function by a minimal collection of maximal intervals, that is generated by the GISI algorithm; *ii* - constructing the dynamical envelope; *iii* - projecting the learned system in the constrained space through the operation of infimum and supremum.

This kind of technique may be useful for the hybrid design of FLDS. The researcher uses his knowledge about the target FLDS to design heuristically (possibly interacting with a simulator [2]) an hypothesis and, then, relax his model constructing a dynamical envelope around it. The identification from samples of the target FLDS, under this envelope constraint, permits the integration of prior knowledge with data information. This kind of technique could be used in several applied areas, for instance, modeling of Biological systems such as genetic or neural networks.

## References

[1] M. Anthony and N. Biggs. *Computational Learning Theory - An Introduction.* Cambridge University Press, 1992.

[2] H. A. Armelin, J. Barrera, E. R. Dougherty, J. E. Ferreira, M. D. Gubitoso, N. S. T. Hirata, and E. J. Neves. Simulator for Gene Expression Networks. In *Microarrays: Optical Technologies and Informatics, Proceedings of SPIE Vol. 4266*, pages 248–259, 2001.

[3] G. J. F. Banon and J. Barrera. Decomposition of Mappings between Complete Lattices by Mathematical Morphology, Part I. General Lattices. *Signal Processing*, 30:299–327, 1993.

[4] J. Barrera, E. R. Dougherty, and N. S. Tomita. Automatic Programming of Binary Morphological Machines by Design of Statistically Optimal Operators in the Context of Computational Learning Theory. *Electronic Imaging*, 6(1):54–67, January 1997.

[5] E. Dougherty, J. Barrera, M. Gubitoso, and N. S. T. Hirata. Dynamical systems with morphological transitions. In *Proceedings of NSIP'01*, 2001.

[6] R. Hirata Jr., N. S. T. Hirata, J. Barrera, and E. R. Dougherty. Two Extensions to the ISI Algorithm. *In preparation*, 2002.