BAYESIAN ESTIMATION OF DISCRETE CHAOTIC SIGNALS BY MCMC

David Luengo, Carlos Pantaleón, Ignacio Santamaría
DICOM, ETSII y Telecomunicación, Universidad de Cantabria.
Avda Los Castros s.n., 39005 Santander, Spain.
Tel: +34 942 201388; fax: +34 942 201488
e-mail:david@gtas.dicom.unican.es *

ABSTRACT

This paper considers Markov Chain Monte Carlo (MCMC) methods for the estimation in Additive White Gaussian Noise (AWGN) of discrete chaotic signals generated iterating any unimodal map. In particular, the Metropolis-Hastings (MH) algorithm is applied to the estimation of signals generated by iteration of the logistic map. Using this technique, Bayesian Minimum Mean Square Error (MS) and Maximum a Posteriori (MAP) estimators have been developed for any unimodal map. Computer simulations show that the proposed algorithms attain the Cramer-Rao Lower Bound (CRLB), and outperform the existing alternatives.

1 INTRODUCTION

Chaotic signals have received much attention during the last years. We will concentrate on discrete chaotic signals, i.e. signals generated iterating a system of difference equations (nonlinear map). These signals are potentially attractive for a wide variety of signal processing applications ranging from time-series modeling to communications. However, classical signal processing techniques are not adequate for this class of signals, which possess properties typical of random signals, in spite of being deterministic. Consequently, there is a need for robust and efficient detection and estimation algorithms for this kind of signals in noise.

The Maximum-Likelihood (ML) estimator of signals generated by any Piecewise Linear (PWL) map has been obtained in [1]. However, it has a high computational cost, that may not be reduced except for certain cases, such as the tent-map, for which an efficient algorithm has been developed in [2]. Bayesian estimators have also been proposed in [3] for signals generated by the tent-map, and extended to any PWL map in [4]. However, these methods can only be applied to PWL maps, and signal estimation alternatives proposed for unimodal maps are in general suboptimal. In [5] a dynamical programming ML estimator, as well as a method based on the itinerary of the signal called "halving method", are proposed. In [6] a method based on the connection between the symbolic sequence associated to a particular signal and its initial condition is presented. Bayesian estimators for signals generated by any unimodal map have not been considered yet.

On the other hand, MCMC techniques are an effective way of generating samples from a complicated probability density function. The most widespread MCMC algorithm is the MH algorithm [7], which proceeds by selecting a candidate point according to a proposal distribution, and accepting it with a certain probability [7]. MCMC based algorithms have been extensively used recently to solve a large variety of problems in communications [8], time-series modeling and prediction [9], and statistical physics and artificial intelligence [7].

In this work, we develop Bayesian estimators based on the MH algorithm for chaotic signals generated iterating any unimodal map. In particular, MS and MAP estimators are developed for signals whose dynamics are governed by the logistic map. The selection of the prior density is based on the invariant density associated with the chaotic sequences. Since we are not able to sample directly from the resulting posterior density, we will make use of the MH algorithm. To avoid the numerical instability characteristic of chaotic signals, an alternative parameterization is considered, and the sequence is generated by backward iteration. The resulting MAP and MS estimators outperform the existing suboptimal alternatives, are asymptotically unbiased, and attain the CRLB for a high Signal to Noise Ratio (SNR).

2 THE LOGISTIC MAP

The signals that we consider in this work are generated according to

$$x[n+1] = F(x[n]), \tag{1}$$

where $F(\cdot)$ is a nonlinear noninvertible map, that we will assume known and unimodal. In this work we will concentrate on the so called logistic map, which is defined as

$$F(x) = \lambda x(1-x), \tag{2}$$

for some bifurcation parameter $1 < \lambda \leq 4$. This map shows chaotic behaviour for $\lambda > 3.5699456$ [10], with a unique attractor for each value of λ that is reached by most initial conditions. However, in contrast with the family of tent-maps, there are regular windows for certain values of λ inside the chaotic region [10]. Although the logistic map is noninvertible, as it is unimodal, it has only two inverse images, which can be easily obtained from

$$x[n] = F^{(-1)}(x[n+1]) = 0.5\left(1 \pm \sqrt{1 - 4x[n+1]/\lambda}\right). \tag{3}$$

We may divide the phase space of the logistic map into two non-overlapping regions $E_1 = [0, 0.5)$ and $E_2 = [0.5, 1]$, and associate a symbol $s[n]$ to each $x[n]$ according to

$$s[n] = \text{sign}(x[n] - 0.5). \tag{4}$$

Now we can define the sequence $\mathbf{s} = \{s[0] \ldots s[N-1]\}$ associated with a length $N+1$ chaotic signal, which will be

* This work has been partially supported by the European Commission and the Spanish Government under Grant 1FD97-1066-C02-01 and TIC 2001-0751-C04-03.

its itinerary, and can be considered a symbolic coding of the chaotic signal. For $\lambda = 4$, which will be the parameter value considered throughout the work, it is straightforward to express (3) as a function of $s[n]$ using (4)

$$x[n] = F^{(-1)}_{s[n]}(x[n+1]) = 0.5\left(1 + s[n](1 - x[n+1])^{1/2}\right),$$

where the subindex $s[n]$ in $F^{(-1)}(\cdot)$ stresses the fact that backward iteration requires knowledge of the sign of the previous sample.

3 BAYESIAN ESTIMATION

3.1 Problem Statement

The signal model we are considering is

$$y[n] = x[n] + w[n] \qquad n = 0, ..., N.$$

The chaotic signal that we wish to estimate $x[n]$ is generated iterating (2) according to (1) for an unknown initial condition $x[0] \in (0, 1)$, and $w[n]$ is a stationary AWGN process with zero mean, and variance σ^2.

3.2 Prior Density

To develop Bayesian estimators, we need to define the prior density for the whole sequence $\mathbf{x} = \{x[0] \; ... \; x[N]\}$. But first we require the prior density for a single point of the sequence. The natural choice is to assign the invariant density associated with the signals generated by the logistic map. When $\lambda = 4$, it is possible to obtain a closed-form expression for the invariant density of the sequence using the Frobenius-Perron operator, which yields [10]

$$p(x) = \frac{1}{\pi\sqrt{x(1-x)}} \qquad 0 < x < 1. \tag{5}$$

Thus, as the whole sequence is completely defined for a given value of $x[0]$, the prior density of $\mathbf{x}$ for $\lambda = 4$ will be

$$p(\mathbf{x}) = \frac{\prod_{n=1}^{N} \delta\left(x[n] - F^{(n)}(x[0])\right)}{\pi\sqrt{x[0](1-x[0])}} \qquad 0 < x[n] < 1. \tag{6}$$

Where $F^{(n)}(\cdot)$ denotes de n-fold functional composition of F, and δ is Dirac's delta. For all the other values of λ in the chaotic region no closed-form expressions are available. In these cases a staircase approximation of the invariant density, as in [3], will be considered.

3.3 Posterior Density

Since our observations $\mathbf{y} = \{y[0] \; ... \; y[N]\}$ are a collection of independent Gaussian random variables with equal variance, their conditional density may be expressed as

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{(N+1)/2}} \exp\left(-\frac{J(\mathbf{x})}{2\sigma^2}\right), \tag{7}$$

where

$$J(\mathbf{x}) = \sum_{n=0}^{N} (y[n] - x[n])^2 = \|\mathbf{y} - \mathbf{x}\|_2^2.$$

Using (6) and (7), and applying the Bayes rule, the posterior density of the sequence $\mathbf{x}$ becomes

$$p(\mathbf{x}|\mathbf{y}) = K \frac{\prod_{n=1}^{N} \delta\left(x[n] - F^{(n)}(x[0])\right)}{\sqrt{x[0](1-x[0])}} \exp\left(-\frac{J(\mathbf{x})}{2\sigma^2}\right), \tag{8}$$

where K is a normalization constant. The dependence on $x[1], ..., x[N]$ can be eliminated by integration, thus obtaining the marginal density

$$p(x[0]|\mathbf{y}) = \frac{K}{\sqrt{x[0](1-x[0])}} \exp\left(-\frac{J(\mathbf{x}_0)}{2\sigma^2}\right), \tag{9}$$

where $\mathbf{x}_0 = \{x[0] \; F(x[0]) \; ... \; F^{(n)}(x[0])\}$ is the whole sequence expressed as a function of $x[0]$ using (1) and (2).

In this case it is not possible to obtain Bayesian estimators directly from (8) or (9), since we do not have a closed-form expression for $F^{(n)}(\cdot)$, as in the case of PWL maps [1]. A gradient-descent approximation is not possible either, because J is not quadratic in the initial condition, as in [3], but a polynomial of order 2^{N+1}. Moreover, the estimation of the chaotic signal using (9) will suffer from the propagation of errors and numerical instability characteristic of chaotic systems. Therefore, the complementary problem will be considered: estimating the last point of the sequence and the itinerary of the signal. For this purpose, we will use an hybrid parameterization of the signal as a function of the itinerary and $x[N]$:

$$\boldsymbol{\theta} = \{s[0] \; ... \; s[N-1] \; x[N]\}.$$

Now $x[N]$ and the itinerary may be expressed as $x[N] = \boldsymbol{\theta}(N+1)$ and $\mathbf{s} = \boldsymbol{\theta}(1\colon N)$ respectively. And, the whole sequence becomes $\mathbf{x}_{\boldsymbol{\theta}} = \{F^{(-N)}_{\boldsymbol{\theta}(1:N)}(x[N]) \; ... \; F^{(-1)}_{\boldsymbol{\theta}(N)}(x[N]) \; x[N]\}$. The prior density in this case has an expression similar to (6), but substituting the information on the $x[n]$ by the itinerary:

$$p(\boldsymbol{\theta}) = \frac{\sum_{n=1}^{2^N} p_n \delta\left(\mathbf{s} - \mathbf{s}_n\right)}{\pi\sqrt{x[N](1-x[N])}}.$$

Being $\mathbf{s}_n$ each of the 2^N possible itineraries, and p_n the probability of each itinerary. Hence, the posterior density now may be expressed as

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{K}{\sqrt{x[N](1-x[N])}} \sum_{n=1}^{2^N} \delta\left(\mathbf{s} - \mathbf{s}_n\right) \exp\left(-\frac{J(\mathbf{x}_{\boldsymbol{\theta}_n})}{2\sigma^2}\right), \tag{10}$$

where $\mathbf{x}_{\boldsymbol{\theta}_n}$ is the sequence obtained iterating backwards from $x[N]$ usign $\mathbf{s}_n$, and, since the itineraries are equiprobable, $p_n = p$, and may be incorporated into K.

3.4 MAP and MS Bayesian Estimators

Two Bayesian estimators will be considered: MAP and MS. The MAP estimator of the sequence is given by the value of $\mathbf{x}$ that maximizes (8). However, since the MAP estimate is always a valid sequence $\mathbf{x}$ (one that agrees with the signal model given by (1) and (2)), it is completely defined by the initial condition $x[0]$, and may be obtained alternatively as the value of $x[0]$ which maximizes (9), or the value of $\boldsymbol{\theta}$ which maximizes (10):

$$\begin{aligned} \hat{\mathbf{x}}_{\text{MAP}}(0) &= \hat{x}_{\text{MAP}}[0] = \arg\max_{x[0]}\{p(x[0]|\mathbf{y})\} = \\ &= F^{(-N)}_{\hat{\boldsymbol{\theta}}_{\text{MAP}}(1:N)}(\hat{\boldsymbol{\theta}}_{\text{MAP}}(N+1)). \end{aligned} \tag{11}$$

The MS estimate of the sequence is the mean of the posterior density given by (8). In this case the MS estimate of the sequence will not be a valid sequence in general, i.e. the MS estimate of the sequence $\hat{\mathbf{x}}_{\text{MS}}$ obtained integrating (8) will not be a sequence generated iterating (2) according to (1). Therefore, (11) is not valid for the MS estimator, and the MS estimates obtained using (9) and (10) will be projections of $\hat{\mathbf{x}}_{\mathrm{MS}}$ over the subset of valid sequences.

4 MCMC BAYESIAN ESTIMATORS

In this paper we consider the use of the MH algorithm [7] to obtain samples from the posterior density $p(\boldsymbol{\theta}|\mathbf{y})$. As a starting point for the algorithm we will use the Hard-Censoring (H-C) estimate of $\boldsymbol{\theta}$, i.e. $\boldsymbol{\theta}_0 = \{\mathrm{sign}(y[0]-0.5)\ ...\ \mathrm{sign}(y[N-1]-0.5)\ y[N]\}$. The steps for the k-th iteration of the algorithm are:

1. Generate a random variable x according to the invariant distribution of the tent-map, uniform in [0,1] [10].
2. Use the equivalence relationship $h(x)$ between the tent-map and the logistic map [10]
$$h(x) = \sin^2(\pi x/2)$$
to obtain a sample $x_k[N] = h(x)$ according to the invariant distribution of the logistic map.
3. Construct a candidate $\boldsymbol{\phi}_k = \{\boldsymbol{\theta}_k(1\colon N)\ x_k[N]\}$
4. Accept the new sequence, thus updating the current state, according to the acceptance function
$$A(\boldsymbol{\theta}_k, \boldsymbol{\phi}_k) = \min(1, Q(\boldsymbol{\theta}_k, \boldsymbol{\phi}_k)),$$
being $\boldsymbol{\theta}_k = \{\boldsymbol{\theta}_k(1\colon N)\ x_k[N]\}$ the current state, and $Q(\boldsymbol{\theta}_k, \boldsymbol{\phi}_k)$ the relationship between $p(\boldsymbol{\phi}_k|\mathbf{y})$ and $p(\boldsymbol{\theta}_k|\mathbf{y})$
$$Q(\boldsymbol{\theta}_k, \boldsymbol{\phi}_k) = \left(\frac{\boldsymbol{\theta}_k(N+1)(1-\boldsymbol{\theta}_k(N+1))}{\boldsymbol{\phi}_k(N+1)(1-\boldsymbol{\phi}_k(N+1))}\right)^{1/2} \exp\left(\frac{J(\mathbf{x}_{\boldsymbol{\theta}_k}) - J(\mathbf{x}_{\boldsymbol{\phi}_k})}{2\sigma^2}\right).$$
5. Construct a new candidate $\boldsymbol{\phi}_k^{(0)} = \{-\boldsymbol{\theta}_k^*(1)\ \boldsymbol{\theta}_k^*(2\colon N+1)\}$, where $\boldsymbol{\theta}_k^*$ is the updated current state.
6. Accept the new sequence, thus updating the current state, according to the acceptance function
$$A(\boldsymbol{\theta}_k^*, \boldsymbol{\phi}_k^{(0)}) = \min(1, Q(\boldsymbol{\theta}_k^*, \boldsymbol{\phi}_k^{(0)})), \qquad (12)$$
where
$$Q(\boldsymbol{\theta}_k^*, \boldsymbol{\phi}_k^{(0)}) = \exp\left(\frac{J(\mathbf{x}_{\boldsymbol{\theta}_k^*}) - J(\mathbf{x}_{\boldsymbol{\phi}_k^{(0)}})}{2\sigma^2}\right).$$
7. Repeat steps 5 and 6 for $s[1], ..., s[N-1]$, i.e. for each $s[n]$ construct a candidate $\boldsymbol{\phi}_k^{(n)} = [\boldsymbol{\theta}_k^*(1\colon n)\ \ -\boldsymbol{\theta}_k^*(n+1)\ \ \boldsymbol{\theta}_k^*(n+2\colon N+1)]$, and accept the new candidate, thus updating the current state, according to (12).
8. Store the $N+1$ sequences obtained and go back to step 1 to start iteration $k+1$.

To allow the distribution of the Markov chain to converge to $p(\boldsymbol{\theta}|\mathbf{y})$ we establish a "burning period" [7], i.e. the first iterations of the algorithm will not be considered for the Bayesian estimation. From the samples obtained using the previous algorithm it is straightforward to obtain the MAP estimate, as the sequence $\boldsymbol{\theta}_k$ which maximizes (10), and, through backward iteration obtain $\hat{x}_{\mathrm{MAP}}[0]$.

The MS estimate of the sequence will not be a valid sequence in general. However, the MS estimate for any point of the sequence can be easily obtained using Monte Carlo integration:

$$\hat{x}_{\mathrm{MS}}[n] = \frac{1}{M}\sum_{i=1}^{M} F_{\boldsymbol{\theta}_i(n:N)}^{(-(N-n))}(\boldsymbol{\theta}_i(N+1)), \qquad (13)$$

where M is the number of samples obtained ($N+1$ multiplied by the number of iterations). Using (13) to estimate $x[0]$ and iterating forward, we obtain an estimator for the whole sequence that we will call MH-MS1. $\hat{\mathbf{x}}_{\mathrm{MH\text{-}MS1}}$ is a projection of $\hat{\mathbf{x}}_{\mathrm{MS}}$ over the subset of valid sequences, that shares the initial condition with $\hat{\mathbf{x}}_{\mathrm{MS}}$. This can be viewed as the best MS estimator if we are only interested in the Mean Square Error (MSE) of $\hat{x}[0]$. However, if we are interested in minimizing the MSE of the whole sequence, $\hat{\mathbf{x}}_{\mathrm{MH\text{-}MS1}}$ is not the best projection. A better estimator, that we will call MH-MS2, is obtained using $\hat{\mathbf{x}}_{\mathrm{MS}}[N]$, and iterating backwards using the most likely itinerary to generate the whole sequence $\hat{\mathbf{x}}_{\mathrm{MH\text{-}MS2}}$. $\hat{\mathbf{x}}_{\mathrm{MH\text{-}MS2}}$ is a projection of $\hat{\mathbf{x}}_{\mathrm{MS}}$ over the subset of valid sequences, that shares the itinerary and ending point of the sequence with $\hat{\mathbf{x}}_{\mathrm{MS}}$.

5 MCMC SIMULATION RESULTS

In this section we analyze the performance of the three estimators considered. First, we study their behaviour for short sequences, namely with $N = 5$. Fig. 1 shows an example of the performance of the three estimators for $x[0] = 0.55$ and 1000 simulations. We use the first 5000 iterations as burning process, and the next 40000 iterations to evaluate the MSE of the initial condition $x[0]$. Since for each iteration $N+1$ candidates are considered, a total of 30000 and 240000 sequences are considered for the burning process and estimation of the MSE respectively. All the MCMC approaches behave similarly, attain the CRLB at approximately 35 dB, and improve considerably the performance of the H-C estimator.

To assess the actual performance of the Bayesian estimators developed, we have conducted a simulation for 1000 initial conditions distributed according to (5). For each initial condition 1000 experiments have been performed, with a burning process of 250 iterations, and 1000 iterations to obtain the MSE. The results for the MSE of $x[0]$ are shown in Table 1, and for the MSE of the whole sequence in Table 2. In both cases the MCMC based estimators outperform the H-C estimator considerably.

Finally, the behaviour of the proposed MCMC algorithms for long registers is investigated. Fig. 2 shows a typical curve for $N = 29$ and $x[0] = 0.688$. In this case the MSE of the whole sequence must be considered. Therefore, only the MH-MS2 algorithm, which concentrates on the whole sequence rather than just on $x[0]$, is shown. Although the MH-MS2 algorithm outperforms the H-C estimator, the gain is not as noticeable as in the case of short registers, while the computational load in this case is much higher.

6 CONCLUSIONS

In this work we propose an alternative to obtain MS and MAP Bayesian estimators using MCMC techniques for any chaotic signal generated by iterating a unimodal map, and observed in AWGN. In particular, the MH algorithm is applied to the estimation of signals generated by the logistic map. First, we construct the posterior density of the sequence using the invariant density associated with the chaotic signals as prior density. Since we are not able to sample directly from the resulting density, we use the

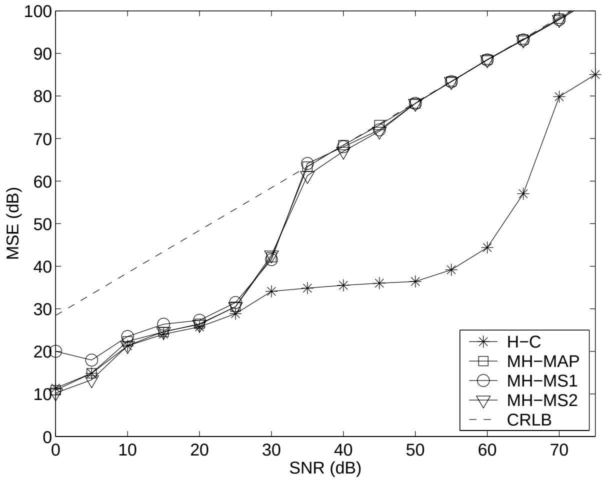

Figure 1: MSE(dB) of x[0] for $N = 5$, $x[0] = 0.55$, $N_{\text{test}} = 1000$, $N_{\text{burning}} = 5000$, and $N_{\text{iter MCMC}} = 40000$.

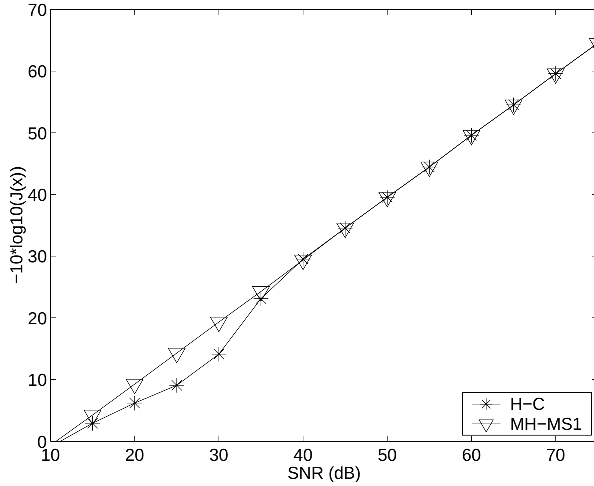

Figure 2: MSE(dB) of the sequence for $N = 29$, $x[0] = 0.688$, $N_{\text{test}} = 1000$, $N_{\text{burning}} = 1000$, and $N_{\text{iter MCMC}} = 10000$.

	MSE (dB)				
		MCMC			
SNR (dB)	H-C	MAP	MS1	MS2	CRLB
10	17.5	17.9	18.8	17.5	41.7
15	25.1	25.4	26.0	25.0	46.7
20	32.9	33.4	33.9	32.9	51.7
25	40.2	42.0	42.4	41.1	56.7
30	44.9	50.4	50.5	49.1	61.7
35	47.5	59.5	59.2	57.6	66.7
40	49.1	67.1	66.8	65.2	71.7
60	55.3	90.6	89.4	89.3	91.7

Table 1: Average CRLB and MSE of $x[0]$ for the four estimators considered, $N = 5$ and 1000 initial conditions.

	MSE (dB)			
		MCMC		
SNR (dB)	H-C	MAP	MS1	MS2
10	5.4	7.9	5.3	7.2
15	8.5	12.5	10.4	11.8
20	11.1	17.4	15.5	16.7
25	13.2	22.3	20.5	21.8
30	14.8	27.3	25.6	26.9
35	16.2	32.3	30.6	32.0
40	17.8	37.2	35.7	37.0
60	23.9	57.0	55.9	56.7

Table 2: MSE of the whole sequence for the four estimators considered, $N = 5$ and 1000 initial conditions.

MH algorithm to obtain samples of $x[N]$ and the itinerary. Then, through backward iteration, we obtain samples of the whole chaotic sequence, which allow us to calculate MS and MAP Bayesian estimators. Using computer simulations the MCMC Bayesian estimators are shown to outperform the existing suboptimal alternatives. The main disadvantage of these techniques is the high computational cost associated to MCMC algorithms. Further research lines include applying this technique to other maps, and studying ways to accelerate the convergence rate.

References

[1] C. Pantaleón, D. Luengo, I. Santamaria, "Optimal Estimation of Chaotic Signals Generated by Piecewise-Linear Maps", *IEEE Signal Processing Letters*, vol. 7, no. 8, pp. 235-237, Aug. 2000.

[2] H. C. Papadopoulos, G. W. Wornell, "Maximum Likelihood Estimation of a Class of Chaotic Signals", *IEEE Trans. Info. Theory*, vol. 41, no. 1, pp.312-317, 1995.

[3] C. Pantaleón, D. Luengo, I. Santamaria, "Bayesian Estimation of a Class of Chaotic Signals", *Proc. ICASSP 2000*, vol. 1, pp. 197-200, Istanbul, Turkey, June 2000.

[4] C. Pantaleón, L. Vielva, D. Luengo, I. Santamaria, "Bayesian Estimation of Chaotic Signals Generated by Piecewise-Lineal Maps", *Signal Processing*, submitted.

[5] S. M. Kay, "Methods for Chaotic Signal Estimation", *IEEE Trans. Signal Processing*, vol. 43, no. 8, pp. 2013-2016, 1995.

[6] L. Cong, W. Xiafou, S. Songgeng, "A General Efficient Method for Chaotic Signal Estimation", *IEEE Trans. Signal Processing*, vol. 47, no. 5, pp. 1424- 1428, 1999.

[7] R. M. Neal, "Probabilistic Inference Using Markov Chain Monte Carlo Methods", *Technical Report CRG-TR-93-1*, Dep. of Computer Science, University of Toronto, 25 September 1993.

[8] S. Lesage, J. Y. Tourneret, P. M. Djuric, "Classification of Digital Modulations by MCMC Sampling", *Proc. ICASSP 2001*, vol. 4, pp. 2553-2556, Salt Lake City, Utah, May 7-11, 2001.

[9] Y. Nakada, T. Kurihara, T. Matsumoto, "Bayesian MCMC Nonlinear Time Series Prediction", *Proc. ICASSP 2001*, vol. 6, pp. 3509-3512, Salt Lake City, Utah, May 7-11, 2001.

[10] H. O. Peitgen, H. Jurgens, D. Saupe "Chaos and Fractals: New Frontiers of Science" *Springer-Verlag*, 1992.