International Journal of Engineering & Technology
CENTRE OF PROFESSIONAL RESEARCH PUBLICATIONS
ISSN 2049-3444

International Journal of Engineering and Technology Volume 3 No. 5, May, 2013

A Particle Filter Based Method for Evaluation of Information Gap between Dynamical Systems

Zhen Zhen[1] Jun Young Lee[2], Hyun Jun Kim[3], Abdus Saboor[4]
[1]Mingde College, Guizhou University, China
[2,3] Department of Computer Science, Seoul National University, Korea
[4] Department of Mechanical Engineering, Stamford University Bangladesh, Bangladesh

ABSTRACT

This paper presents a new result on evaluating the difference between two dynamical systems. Based on the idea of information theoretic gap developed in [12], a new numerical method is developed to compute the Kullback-Leibler (K-L) rate pseudo metric between dynamics systems. Our method is based on SIR particle filter and multimodal Gaussian representation of particles is used for the computation of K-L divergence. This proposed method is relatively easy to implement and is capable of handling nonlinear systems. Numerical experiments are provided to show the efficacy of this approach.

Keywords: *Kullback-Leibler Divergence, Dynamical System, Particle Filter, Information Theory*

1. INTRODUCTION

With the advances of communication technologies, researchers become more and more interested in how information can be transmitted or stored from the perspective of end user [1, 2]. However from the perspective of automation and control system it is as important to consider information in the classic sense of reduction of uncertainty. How much information is generated when the dynamical system enters a particular state by either external disturbance or internal interactions among its components? And to what extent is the information transmitted between subsystems affect the behavior of a control system? The latter problem is of significant practical importance as networked control system becomes a popular topic in research and industrial applications. A large number of publications on the stability and performance of networked control system can be found in the literature [3, 4, 5, 11]. More generally, the performance limitation for feedback control system was also studied from the perspective of information processing. The results along this thread of research imply that without sufficient feedback information, the performance of control system can be degraded [6, 9, 10, 13, 14]. These works motivated the study of information processing within dynamical systems. The objective, broadly speaking, is to understand how uncertainties propagate in dynamical systems, how these information constraints can be used to characterize different dynamical systems, and how these properties decide the behavior of the overall system formed by connecting multiple dynamical systems to each other.

To achieve this objective, it is important to first understand the difference between dynamical systems, from information perspective. In a recent work [12], the authors studied the propagation of uncertainties in different dynamical systems, and further quantified the difference between systems using Kullback-Leibler divergence rate. Information distance or divergence measures are of key importance in a number of theoretical and applied statistical inference and data processing problems, such as estimation, detection, classification, compression, recognition, etc. Kullback-Leibler divergence rate is one of these measures derived from Kullback-Leibler divergence, it is also closely related to maximum likelihood and Bayesian identification procedures. This metric is not different from that used in aggregation of Markov and Hidden Markov models by the same authors [16]. However unlike in Markov and Hidden Markov case, the metric in [12, 17] cannot be easily computed because it is defined for general nonlinear dynamical systems. In order to compute this metric, the posterior distributions of system variables should be computed and compared first. For general nonlinear systems this is very difficult.

In this work we propose a particle filter based framework to approximately compute the Kullback-Leibler divergence rate in [12]. The idea is to use particles as empirical representations of the real posterior distribution. One technical difficulty here is that Kullback-Leibler divergence cannot be directly computed between two distributions in the form of particles, we proposed to construct a multimodal Gaussian distribution from the particle representation so that a continuous approximation of arbitrary posterior distribution can be obtained. The Kullback-Leibler divergence between two posterior distributions can therefore be approximated using the Kullback-Leibler divergence between two constructed multimodal Gaussian distributions. Using the proposed method, Kullback-Leibler divergence rate between two nonlinear dynamical systems can be approximately computed. In section 4, the numerical result for an illustrative example is presented.

2. KULLBACK-LEIBLER RATE METRIC

To make this paper self-contained, in this section we introduce the concept of Kullback-Leibler rate metric based on [12]. In information theory, the Kullback–Leibler divergence (also known as information divergence, information gain, relative entropy) is a non-symmetric measure of the difference between two probability distributions P and Q. Specifically, the Kullback-Leibler divergence of Q from P, denoted as $D(P||Q)$, is a measure of the information lost when Q is used to approximate P [18]: Kullback-Leibler divergence measures the expected number of extra bits required to code samples from P when using a code based on Q, rather than using a code based on P. Typically P represents the "true" distribution of data, observations, or a theoretical distribution. The distribution Q typically represents an approximation of distribution P.

For discrete probability mass distributions P and Q, Kullback-Leibler divergence is defined as:

$$D_{KL}(P||Q) = \sum_i \ln(\frac{P(i)}{Q(i)})P(i) \qquad (1)$$

If p(x) and q(x) are distributions defined for continuous random variables, the Kullback-Leibler divergence is

$$D_{KL}(p||q) = \int_{-\infty}^{\infty} \ln(\frac{p(x)}{q(x)})p(x)dx \qquad (2)$$

For a dynamical system M1 and an assumed model M2, optimal filtering using these two models will generate two sequences of posterior distributions (Figure.1).

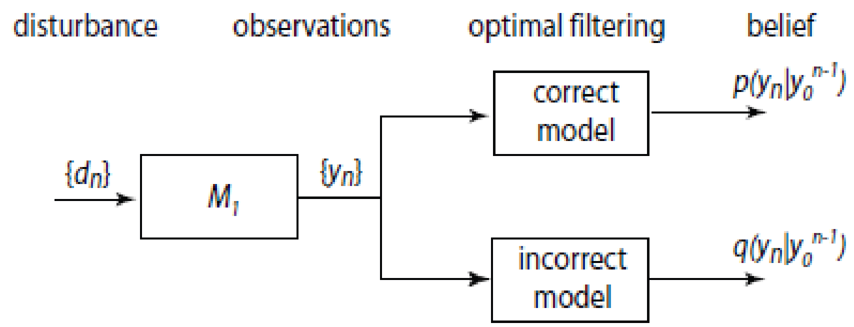

Figure.1 optimal filtering using dynamical system model M1 and M2

The Kullback-Leibler divergence rate is defined as the difference between two entropy rates: one is associated with the posterior distribution sequence obtained by filtering using true model M1, the other is associated with the posterior distribution sequence obtained by filtering using assumed model M2.

$$\Delta H = \hat{H}(y) - H(y) = \lim_{n\to\infty} E_{p(y_0^{n-1})} D_{KL}(p(y_n|y_0^{n-1})||q(y_n|y_0^{n-1})) \qquad (3)$$

The entropy rate obtained with assumed model is:

$$\hat{H}(y) = \lim_{n\to\infty} E_{p(y_0^n)}[-\ln q(y_n|y_0^{n-1})] \qquad (4)$$

The entropy rate obtained with true model is:

$$H(y) = \lim_{n\to\infty} E_{p(y_0^n)}[-\ln p(y_n|y_0^{n-1})] \qquad (5)$$

Here to denote the joint probability distribution of sequence of random variables we used the compact notation: $p(y_0^n) = p(y_0, y_1, \dots, y_{n-1})$. The notation $E_{p(y_0^{n-1})}[.]$ means the expectation taken with respect to the joint probability distribution $p(y_0^{n-1})$.

To compute the Kullback-Leibler divergence rate, we need to sequentially compute the two posterior distributions. This is usually difficult for general nonlinear systems, because in general the Kullback-Leibler divergence between two arbitrary probability distributions is not easy to compute, unless some constraints are added, for example the distributions are known to be Gaussian. In our case the nonlinear dynamics will distort the posterior distributions, so the nice closed form of Kullback-Leibler divergence for Gaussian distributions is not preserve. In the following section we propose to use particle representation of probability distributions in the computation and utilize particle filter to propagate the uncertainty. The Kullback-Leibler divergence can therefore be computed after the particle approximately.

3. COMPUTATION OF KULLBACK-LEIBLER RATE USING PARTICLE FILTER

Particle filter [7, 8, 15, 19] is widely used in engineering research and practice, especially in control engineering, computer vision and statistics. Its advantage over classical Kalman filter is that particle filter can be used not only for linear systems but also for nonlinear systems. Figure.2 shows the performance of particle filter in a very simple target tracking example. The solid blue curve represents the true target position, and the dotted red curve is the estimated trajectory of target generated by particle filter.

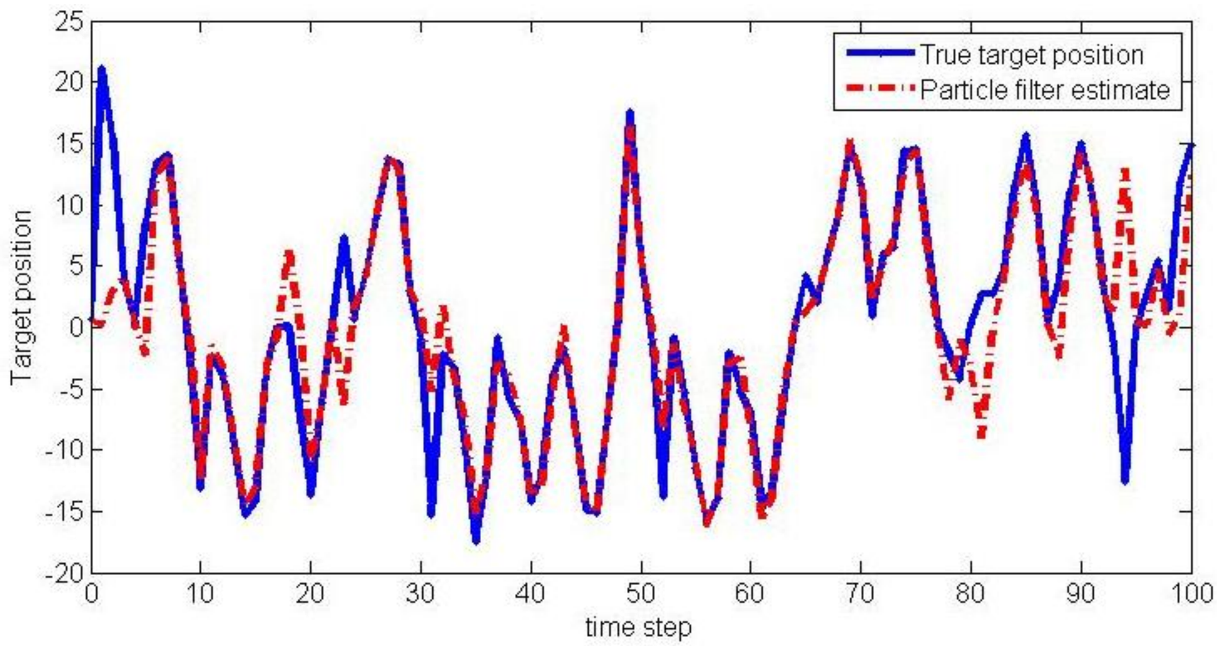

Figure 2.Using particle filter for position tracking

For a dynamical system

$$x_k = g(x_{k-1}) + \omega_k,\ \ y_k = h(x_k) + v_k \qquad (6)$$

Particle filter can generate the posterior distribution sequence for latent states, in the form of particle approximations. In Sequential Importance Resampling algorithm, the filtering distribution $p(x_k|y_0, \ldots, y_k)$ is approximated by a weighted set of P particles

$$\{(\omega_k^{(L)}, x_k^{(L)}): L = 1,2, \ldots, P\} \qquad (7)$$

In (1) the Kullback Leibler divergence should be computed between two particle representations. However the two particle representations for two posterior distributions are of discrete form, and the particles are located at different positions of the space, a direct computation using (1) will lead to infinite numerical value because zero-entry will appear as denominator. A continuous approximation of posterior distributions should be formulated instead so that (2) can be used to compute the Kullback Leibler divergence rate. To do this, we propose to replace each particle with Gaussian kernel so that the posterior distribution becomes multimodal Gaussian distribution. By doing so the discrete representation of weighted particles becomes continuous.

4. NUMERICAL RESULTS

In this section we present an illustrative example to demonstrate the proposed method. Consider two discrete time dynamical systems M1 and M2. System M1 is implemented in MATLAB as:

```
x(n+1) = -0.7*sin(x(n)) + sqrt(sigma)*randn;
   z(n) = x(n)/20 + sqrt(sigma)*randn;
```

Here sigma is the variance of Gaussian noise. System M2 is linear, it is implemented as

```
x(n+1) = -0.5*x(n) + sqrt(sigma)*randn;
```

z(n) = x(n)/20 + sqrt(sigma)*randn;

We use SIR particle filter with 30 particles to recursively generate posterior distributions. At each time step, the instant values of Kullback-Leibler divergence between two posterior distributions can be computed (Figure 3).

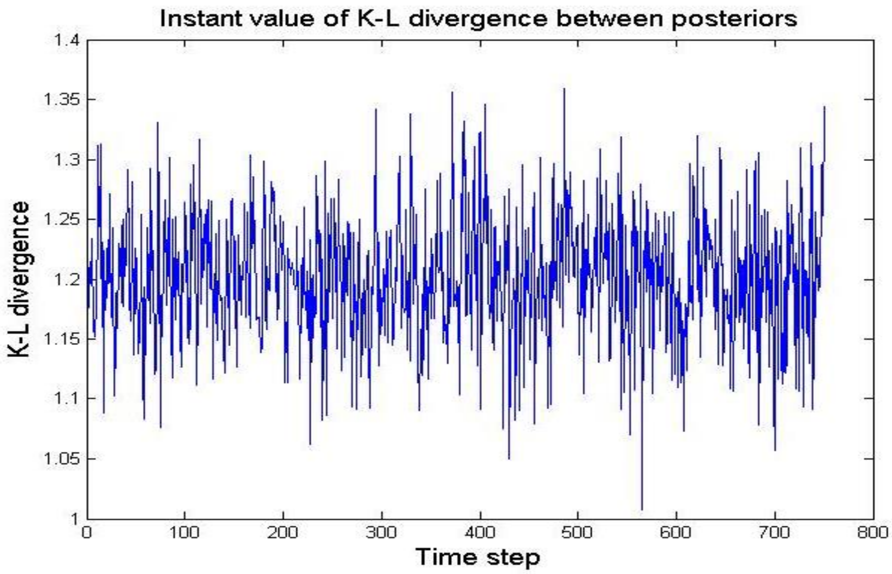

Figure 3. Kullback-Leibler divergence between posterior distributions

The Kullback-Leibler divergence rate is computed from equation (3). We obtain empirical value by running particle filter for 5000 steps and use the average of last 500 points to approximate the asymptotic value. For system M1 and M2, a Kullback-Leibler divergence rate of 1.205 is obtained.

We want to point out that the final numerical value is affected by the number of particles chosen and the variance of Gaussian kernels use in the approximation. The values can be used as a relative quantitative measure of gaps between different dynamical systems.

5. CONCLUSION

In this work we proposed a particle filter based method for computing Kullback-Leibler divergence rate metric in [12]. The advantage of using particle filter is that the nonlinear dynamics can be properly managed. At each time step, multimodal Gaussian mixture distributions are generated from particle representations to facilitate the computation of Kullback-Leibler divergence. Our method can be used to evaluate relative gaps between dynamical systems. Future works include refining this method to reduce its computation time and enhance its accuracy.

REFERENCES

[1] Lerman, K, “Social Information Processing in News Aggregation”, Internet Computing, IEEE. Volume: 11, Issue: 6, Page(s): 16 - 28, 2007.

[2] Noble, J., Davy, S. and Franks, D. W. (2004) “Effects of the topology of social networks on information transmission”, In From Animals to Animats 8: Proceedings of the Eighth international Conference on Simulation of Adaptive Behavior. , MIT Press, 395-404.

[3] Sekhar Tatikonda, S. Mitter, “Control Under Communication Constraints”, IEEE Transactions on Automatic Control, 49, 1056-1068., 2004,

[4] Sekhar Tatikonda, S. Mitter, “Control Over Noisy Channels”, IEEE Transactions on Automatic Control, 49, 1196-1201, 2004,

[5] G. N. Nair & R. J. Evans, “Stabilization with data-rate-limited feedback: tightest attainable bounds”, Systems & Control Letters, vol. 41, no. 1, pp. 49-56, 2000.

[6] Sun Yu, Prashant G. Mehta, “Fundamental performance limitations with Kullback-Leibler control cost,” in Procs. of Conference on Decision & Control, pp.7063-7068, 2010.

[7] M. Sanjeev Arulampalam, Simon Maskell and Neil Gordon, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking", IEEE

Transactions on Signal Processing, vol. 50, pp.174--188, 2002

[8] Doucet, A., De Freitas, N.and Gordon, N.J. "Sequential Monte Carlo Methods in Practice". Springer, 2001.

[9] Martins, Nuno C., Dahleh, Murither A., "Feedback Control in the Presence of Noisy Channels: 'Bode-Like' Fundamental Limitations of Performance", IEEE Transactions on Automatic Control, vol.53, issue.7, pp. 1604 – 1615, 2008

[10] Y. Sun and P. G. Mehta, “Fundamental performance limitations via entropy estimates with hidden Markov models,” in Procs. of Conference on Decision & Control, pp. 3982–3988, 2007.

[11] Nicola Elia: Remote stabilization over fading channels. Systems & Control Letters, vol.54, no.3, pp.237-249, 2005.

[12] Sun Yu, Prashant G. Mehta, “The Kullback-Leibler Rate Pseudo-Metric for Comparing Dynamical Systems”. IEEE Transaction on Automatic Control, vol.55, pp.1585-1598, 2010.

[13] G. Zang, and P. A. Iglesias,“Nonlinear extension of Bode’s integral based on an information theoretic interpretation”, Systems and Control Letters, vol.50, pp.11–29, 2003.

[14] Sun Yu, Prashant G. Mehta, “Bode-Like Fundamental Performance Limitations in Control of Nonlinear Systems”. IEEE Transaction on Automatic Control, vol. 55, pp.1390-1405, 2010.

[15] Doucet, A., De Freitas, N. and Murphy, K. and Russell, S., "Rao–Blackwellised particle filtering for dynamic Bayesian networks", Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence, pp.176–183, 2000.

[16] Kun Deng, Yu Sun, Prashant Mehta and Sean Meyn, "An information-theoretic framework to aggregate a Markov chain", American Control Conference, pp.731-736, 2009.

[17] Sun Yu, Prashant G. Mehta, “The Kullback-Leibler rate metric for comparing dynamical systems”, in Procs. of Conference on Decision & Control, pp.8363-8368, 2009.

[18] T. M. Cover and J. A. Thomas, “Elements of Information Theory”, 3rd Edition, New York: Wiley Interscience, 2005.

[19] Pitt, M.K., Shephard, N., "Filtering Via Simulation: Auxiliary Particle Filters". Journal of the American Statistical Association, 94 (446): 590–591, 1999