

Copula Based Time Delay Estimation for Linear First Order System

Xuguang Wang, Wei Dong* and Yan Zhang

Hebei Engineering Research Center Simulation & Optimized Control for Power Generation, Baoding 071003, China

*Corresponding author

Abstract—Time Delay is common and inevitable in the dynamic system. The accuracy of system identification will not be guaranteed if time delay is not accurately estimated or handled properly. Using the theory of dependency measurement, this paper studies the delay identification of SISO system. Firstly, the copula theory is introduced into time delay identification. Secondly, use dependency to measure the relationship of input and output signals and then estimate the time delay with this relationship. Finally, several first order system experiments are used to verify the proposed theory.

Keyword—time delay estimation(TDE); first order system; copula; dependence; system identification

I. INTRODUCTION

The problem of identification of parameters in system engineering is an interest area of research and has gained increasing significance. As one of the important aspects of parameters identification research, there has been increasing attention to the time delay estimation. Time delay systems are widely used to model concrete systems in engineering sciences, such as biology chemistry, mechanics and so on[1]. Many results have been reported for the purpose of stability analysis, by assuming that the time delay of the studied systems is known.

There are various methods in the literature for estimating time delays, especially in univariate process. Generalized cross correlation(GCC) is the conventional method for finding time delay[2,3], however, in order to obtain an accurate time delay estimate, the GCC method requires a priori[4]. Meanwhile, if the dynamics between signals is considered, a simple model plus time delay is a practical choice for most cases[5,6]. In the last decade, a few new methods for TDE have also been developed, based on different techniques such as interpolation[7,8], neural network[9], linear matrix inequality. There are also some different methods, and these methods also have limitations. Some methods are “model dependent”, which should know the structure of the system before estimating the time delay. Some methods need to know the scope of the parameter. There are still many methods which are very sensitive to noise[10]. So there may remain some limitations in practical application.

This paper presents a new method to identify time delay when the sample data is distorted by noise or the structure of system is unknown. The proposed method is based on the copula theory. We regard the input signals and the output signals are two random variables. The dependence relationship

between random variables is constructed. And the system delay is determined by this relationship.

II. COPULA

A. Basis Theory of Copula

The researchers often study the relationship of random variables by their joint distribution and marginal distribution[11]. The marginal distribution of variables can be obtained by their joint distribution. Meanwhile, when the random variables are independent of each other, the joint distribution can also be calculated by their marginal distribution. The relationship of random variables also contains function relation and dependence relation. Sklar joined their joint distribution and marginal distribution together as a new function, which called copula, when these variables have dependence relation. In probability theory and statistics, a copula is a multivariate probability distribution for which the marginal probability distribution of each variable is uniform. Copulas are used to describe the dependence between random variables. Their name comes from the Latin for "link" or "tie", similar but unrelated to grammatical copulas in linguistics. Copulas have been used widely in quantitative finance to model and minimize tail risk and portfolio optimization applications.

We first give a proposition in order to prompt copula function:

There are two random variables X_1 & X_2 , let $F_1(x)$ & $F_2(x)$ be their distribution functions. Let $Z_{1,2} = F_{1,2}(x)$, then the random variable Z is uniform-distributed in interval $[0,1]$.

Theorem I is proposed according to the above proposition:

(i) In probabilistic terms, is a 2-dimensional copula if C is a joint cumulative distribution function of a 2-dimensional random vector on the unit cube with uniform marginals[12].

(ii) In analytic terms,

$$C : [0,1] \times [0,1] \rightarrow [0,1]$$

is a 2-dimensional copula if

$C(1,v) = C(v,1) = v$, the copula is equal to u if one argument is v and all others 1.

$C(0, v) = C(v, 0) = 0$, the copula is zero if one of the arguments is zero.

$$C(b_1, b_2) - C(a_1, b_2) - C(b_1, a_2) + C(a_1, a_2) \geq 0, \text{ for all } [a_1, b_1] \times [a_2, b_2] \subseteq [0, 1] \times [0, 1]$$

It has been referred that Sklar gives the definition of copula. Sklar's theorem is also regarded as the theoretical basis of copula. Sklar's Theorem states that any multivariate joint distribution can be written in terms of univariate marginal distribution functions and a copula which describes the dependence structure between the variables[13].

Sklar's theorem[8], named after Abe Sklar, provides the theoretical foundation for the application of copulas. Sklar's theorem states that every multivariate cumulative distribution function

$$H(x_1, x_2) = P[X_1 \leq x_1, X_2 \leq x_2] \quad (1)$$

of a two elemental vector (X_1, X_2) can be expressed in terms of its marginals: $F_i(x) = P[X_i \leq x], i = 1, 2$ and a copula C . Indeed:

In case that the multivariate distribution has a density h , and this is available, it holds further that $h(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2)$

where c is the density of the copula.

The theorem also states that, given H , the copula is unique on $Ran(F_1) \times Ran(F_2)$, which is the cartesian product of the ranges of the marginal cdf's. This implies that the copula is unique if the marginals F_1, F_2 are continuous.

The converse is also true: given a copula $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$ and margins F_1, F_2 then $C(F_1(x_1), F_2(x_2))$ defines a 2-dimensional cumulative distribution function.

Sklar's theorem indicates that the joint distribution of random variables can be constructed by two independent steps:

- (i) Marginal distribution of single random variable.
- (ii) Dependence of random variables, which is called copula.

It should be estimated by some methods because the copula between the random variables is usually unknown. There are two classes of methods among the literatures. The first kind of the methods estimate the undetermined parameters based on the principle of maximum likelihood estimate. But this needs to know the marginal distribution of the variables and the type of copula. We can use empirical formula to estimate copula or specify several types of distribution without knowing these prior knowledge and then estimate the parameters by Bayesian method. Finally, use

some criteria to judge the distribution type, such as AIC, BIC, DIC, etc.

Suppose we have observations

$$(X_1^i, X_2^i), i = 1, 2 \quad (2)$$

from a two elemental vector (X_1, X_2) with continuous margins. The corresponding "true" copula observations would be

$$(U_1^i, U_2^i) = (F_1^n(X_1^i), F_2^n(X_2^i)), i = 1, 2 \quad (3)$$

However, the marginal distribution functions F_i are usually not known. Therefore, one can construct pseudo copula observations by using the empirical distribution functions

$$F_k^n(x) = \frac{1}{n} \sum_{i=1}^n 1(X_k^i \leq x) \quad (4)$$

instead. Then, the pseudo copula observations are defined as

$$(\tilde{U}_1^i, \tilde{U}_2^i) = (F_1^n(X_1^i), F_2^n(X_2^i)), i = 1, 2 \quad (5)$$

The corresponding empirical copula is then defined as

$$C^2(u_1, u_2) = \frac{1}{2} \sum_{i=1}^2 (U_1^i \leq u_1, U_2^i \leq u_2) \quad (6)$$

The components of the pseudo copula samples can also be written as $\tilde{U}_k^i = R_k^i / 2$ where R_k^i is the rank of the observation X_k^i :

$$R_k^i = \sum_{j=1}^2 (X_k^j \leq X_k^i) \quad (7)$$

Therefore, the empirical copula can be seen as the empirical distribution of the rank transformed data.

$C : [0, 1] \times [0, 1] \rightarrow [0, 1]$ is the copula of X_1, X_2 , ρ is the dependence measurement of X_1, X_2 :

$$\rho(X_1, X_2) = \alpha \int_{[0,1]^2} |C(V) - \Pi(V)| dV \quad (8)$$

The role of α is to ensure $\rho(X_1, X_2) \in [0, 1]$.

B. Estimate Time Delay Based on Dependence Measurement

X_1, X_2 are two random variables. Suppose X_1 and X_2 are output and input of a SISO system and $X_2[k]$ is the value of X_2 after a delay of kt_0 .

According to the copula of X_1 and X_2 , their marginal distribution can be transformed as :

$$H[X_1 \leq x_1, X_2[k_2] \leq x_2] = H[X_1 \leq x_1 | X_2[k_2] \leq x_2] \cdot H[X_2[k_2] \leq x_2]$$

The formula of delay estimation:

$$\tilde{k}_2 = \arg \max_{k_2} \rho(X_1, X_2[k_2])$$

shows when the time delay

equals $\tilde{k}_2 t_0$, their dependency is maximum. The time delay is the optimal solution of delay estimation. This method is called dependence method in this paper.

III. SIMULATION RESULTS

In this section, we present the simulation results for determining time delay of first order system. For the sake of comparison, different intensity of noise is added to output signal. Some other methods are also mentioned to show the differences between copula. The procedure of the experiment is as follows:

(i) Get the output $Y = \{y_i\}$ from the system $\frac{1}{Ts+1}e^{-\tau s}$ and $\frac{1}{Ts}e^{-\tau s}$ with given input:

$$X = \{x_i\}, x_i \in [0,1]$$

(ii) Calculate the time delay with $\rho(X[k], Y)$

(iii) Add Gaussian noise to input and output signals and then recalculate their dependency.

Example 1. To first order system:

$$\frac{1}{Ts+1} \cdot e^{-\tau s}, \tau = 30, T = 5, 8, 15, 20, \text{ noise free.}$$

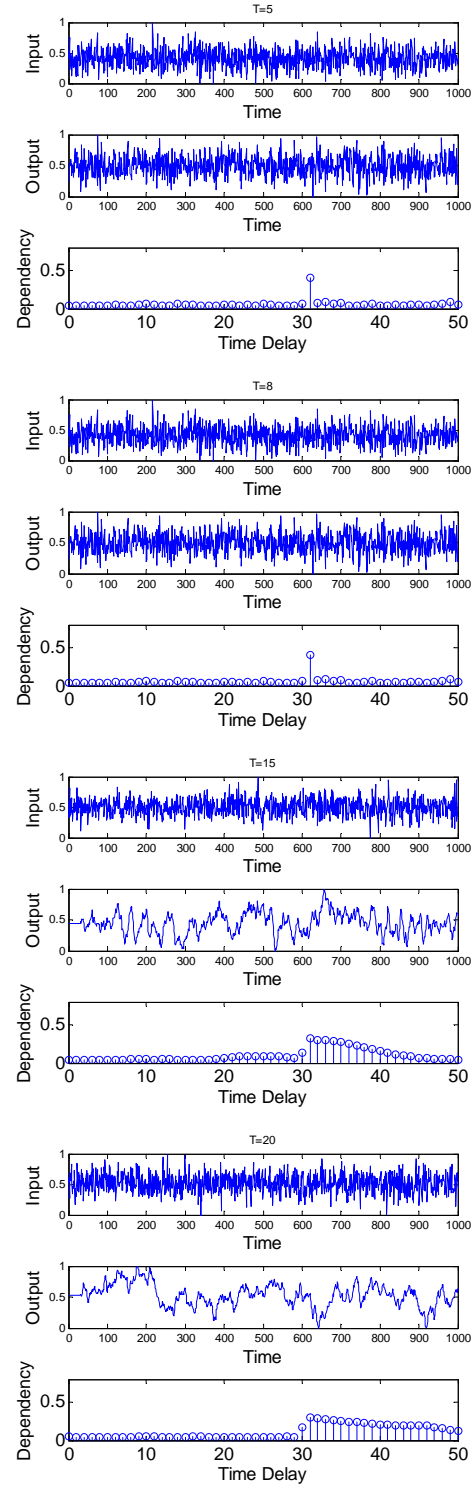


FIGURE I. RESULT OF TDE WITH NO NOISE $\left(\frac{1}{Ts+1} \cdot e^{-\tau s} \right)$
 $T = 5, 8, 15, 20$

Example 2. To first order system:

$$\frac{1}{Ts} \cdot e^{-\tau s}, \tau = 30, T = 5, 8, 15, 20, \text{ noise free.}$$

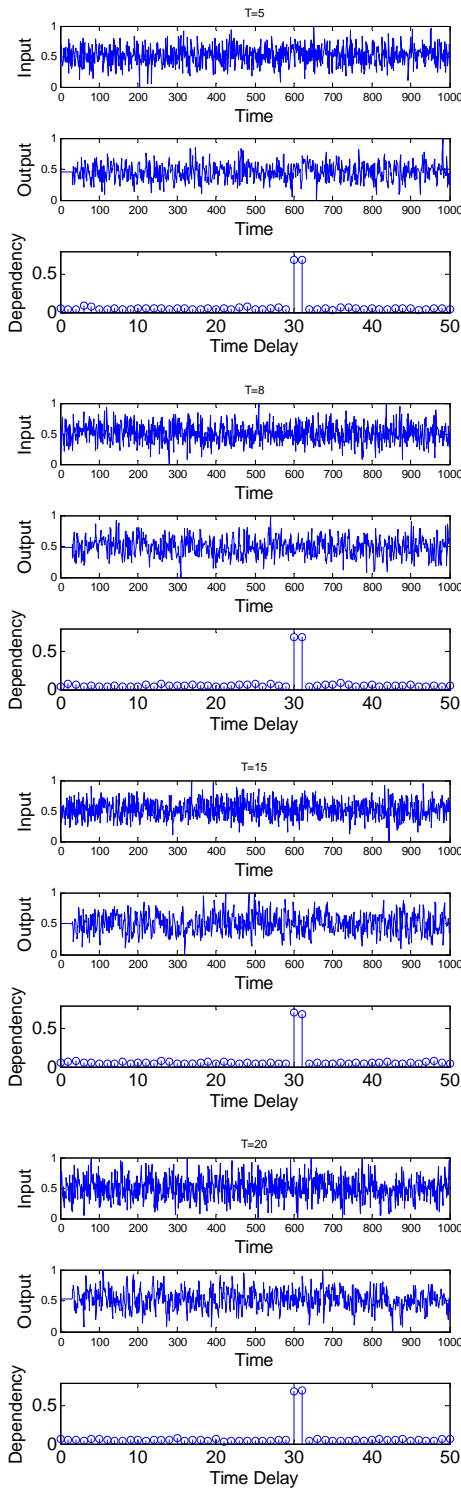


FIGURE II. RESULT OF TDE WITH NO NOISE($\frac{1}{Ts} \cdot e^{-\tau s}$, $T = 5, 8, 15, 20$)

Two conclusions can be drawn from fig I and fig II:

(i) To first order system, the delay estimation obtains a good result, which the computing value is about 30.

(ii) The result is not affected with the increase of T within certain range.

Example 3.

$$\frac{1}{Ts+1} \cdot e^{-\tau s}, \tau = 30, T = 10, \text{ singal-to-noise ratio(SNR)=1,0.5.}$$

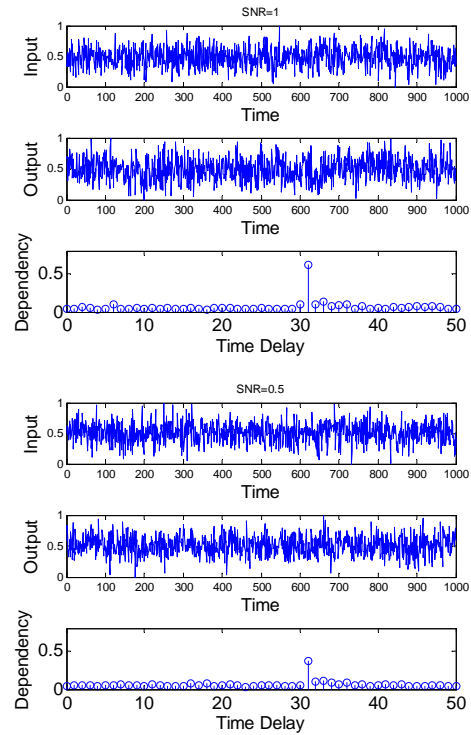


FIGURE III. RESULT OF TDE WITH NOISE($\frac{1}{Ts+1} \cdot e^{-\tau s}$)

Example 4.

$$\frac{1}{Ts} \cdot e^{-\tau s}, \tau = 30, T = 10,$$

singal-to-noise ratio(SNR)=1,0.5.

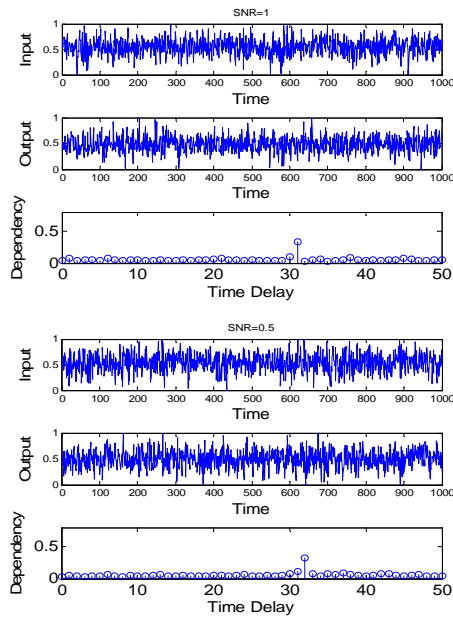


FIGURE IV. RESULT OF TDE WITH NOISE($\frac{1}{Ts} \cdot e^{-Ts}$)

Two conclusions can be drawn from fig.III and fig.IV

(i)To first order system, the delay estimation obtains a good result, which the computing value is about 30.

(ii)The result is not be affected with the increase of noise density within certain range.

Example 5. The comparison between dependence method and other methods

(i)Least square method(LS)

When there is no noise adding to the signals, the calculated result of time delay seems quite good, but when the noise is taken into account, the result get a great error. In addition, the computation cycles is far greater than copula's.

TABLE I. THE COMPARISON BETWEEN LS AND COPULA(NOISE FREE)

Name	Time delay	Noise	Result	MAE
Depende-nce method	30	Noise free	30	0
LS	30	Noise free	30	0

Table I shows the estimation result and mean absolute difference (MAE) of the two methods. The result is the average of 10 simulation data, the same below. It can be seen that when there is no noise added to the signal, LS has the same accuracy with copula.

TABLE II. THE RESULT OF COPULA(NOISE ADDED)

SNR	Time delay	Result	MAE
20	30	30	0
10	30	30.2	0.2
5	30	30.4	0.4
1	30	29.2	0.8
0.5	30	30.8	0.8

TABLE III. THE RESULT OF LS(NOISE ADDED)

SNR	Time delay	Result	MAE
20	30	33.2	3.2
10	30	40	10
5	30	43.8	13.8
1	30	8.8	21.2
0.5	30	3.4	26.6

As we can see in table II and table III, copula has great robustness when the noise is added to the output signals. The time delay estimation result of LS is much worse. The performance is getting worse with the increase of error.

TABLE IV. COMPUTATION CYCLES OF TWO METHODS

Name	Calculating Time
Least square method	Almost 3 min
Dependence method	8s

Table IV shows the comparison of simulating time between the two methods. Obviously the simulating time of LS is much longer copula's.

(ii)Particle swarm optimization(PSO)

Compare to the method proposed in this paper, estimating time delay with PSO needs to know the structure of system, and its result is the approximation of each system parameter. So it sometimes get great error.

IV. CONCLUSION

As a statistical method, dependence method uses marginal distribution and joint distribution when calculates the dependence of random variables. Therefore it can get great results when the variables have certain connection. So we use this method to estimate time delay because of the functional relationship between the input and output signals. This method has the advantages of high computational accuracy, fast computing speed, unaffected by noise. But it has computation complexity, and the future work is to solve this problem.

REFERENCES

- [1] Katz B F, Noisternig M. A comparative study of Interaural Time Delay estimation methods[J]. Journal of the Acoustical Society of America, 2014, 135(135):3530.

- [2] Assous S, Linnett L. High resolution time delay estimation using sliding discrete Fourier transform[J]. Digital Signal Processing, 2012, 22(5):820-827.
- [3] Hayvaci H T, Setlur P, Devroye N, et al. Maximum likelihood time delay estimation and Cramér-Rao bounds for multipath exploitation[C]// Radar Conference. IEEE, 2012:0764-0768.
- [4] Sklar M. Fonctions de Répartition À N Dimensions Et Leurs Marges[J]. Publ.inst.statist.univ. paris, 1960, 8:229-231.
- [5] Johnson N L, Kotz S. Hoeffding, Wassily[M]// Leading Personalities in Statistical Sciences: From the Seventeenth Century to the Present. John Wiley & Sons, Inc. 2012:118-122.
- [6] Isaksson A J, Horch A, Dumont G A. Event-triggered deadtime estimation from closed-loop data[C]// American Control Conference, 2001. Proceedings of the. 2001:3280-3285 vol.4.
- [7] Hou H, Sheng G, Jiang X. Robust Time Delay Estimation Method for Locating UHF Signals of Partial Discharge in Substation[J]. IEEE Transactions on Power Delivery, 2013, 28(3):1960-1968.
- [8] Perdreauville F J, Goodson R E. Identification of Systems Described by Partial Differential Equations[J]. Journal of Fluids Engineering, 1966, 88(2):463-468.
- [9] Peral-Rosado J A D, Lopez-Salcedo J A, Seco-Granados G, et al. Joint maximum likelihood time-delay estimation for LTE positioning in multipath channels[J]. EURASIP Journal on Advances in Signal Processing, 2014, 2014(1):1-13.
- [10] Rao G P, Sivakumar L. Identification of deterministic time-lag systems[J]. IEEE Transactions on Automatic Control, 1976, 21(4):527-529.
- [11] Gawthrop P J. Identification of continuous systems : H. Unbehauen and G. P. Rao[J]. Automatica, 1989, 25(3):476-478.
- [12] Rao V S H, Yadaiah N. Parameter identification of dynamical systems[J]. Chaos Solitons & Fractals, 2005, 23(4):1137-1151.
- [13] Isaksson A J, Horch A, Dumont G A. Event-triggered deadtime estimation from closed-loop data[C]// American Control Conference, 2001. Proceedings of the. 2001:3280-3285 vol.4.
- [14] Ljung L, Glad T. Modeling of dynamic systems[J]. Neurocomputing, 1994, 14(5-6):423-439.