Application of the Optimal Input Design in Shaking Table Experimental System

Jianqiu Chen[1,2] Xinzheng Zhang[1]
1. Department of Automation
University of Guangdong Technology
Guangzhou 510006, P. R. China
cjq_72@126.com

Ping Tan[2] Fulin Zhou[2]
2. Guangdong Key Laboratory of Earthquake Engineering & Applied Technique
University of Guangzhou
Guangzhou 510006, P. R. China

***Abstract*—This paper focuses on the optimal input design question of white noise drive signal for the simulated earthquake shaking table experimental system. The input sequence is white processed by doubly stochastic interchange optimized algorithm which has been improved in this paper, and the operating speed has obviously been expedited. Thus the white noise drive signal can satisfy conditions of the optimal input to a certain extent, and the experimental time won't be so longer.**

***Keywords*—shaking table, PE conditions, optimal input design, system identification**

I. INTRODUCTION

It is desirable to have an accurate process model for the design of a controller. The input signal for the identification experiment determines the nature and accuracy of the system characteristics that are identified and used for control design, which must satisfy PE (Persistence Excitation) conditions. Therefore, the choice of this signal is crucial for the quality and performance of the designed controller. However, there are often constraints on the length of observation time or on the magnitude of disturbance to the system produced by the input signal, which limits the choice of available inputs and makes a careful selection necessary, focusing on the demands prescribed by the intended application of the controller.

Input design problems first arose in statistics and became relevant for engineering in the 1960s. A survey of early work in this area is given in Mehra (1974) [1]. This work focussed mainly on determination of parameters and prediction, while control design was let aside. The design criteria are scalar measures of the information matrix M^{-1}, or the average per data sample information matrix M^{-1}, which were chosen rather ad hoc. Some of these cost functions are:

A-optimality: $tr\overline{M}^{-1}$ (and as a more general case $tr(W\overline{M}^{-1})$), where $W \geq 0$ is a positive semidefinite weighting matrix; this is called *L*-optimality). This measure minimizes the average variance of the parameters.

E-optimality: $\lambda_{max}(\overline{M}^{-1})$ where λ_{max} denotes the maximum eigenvalue.

D-optimality: $\log|\overline{M}^{-1}|$. This measure minimizes the volume of the confidence ellipsoid defined by the covariance matrix of the parameters. An important advantage of *D*-optimality is that it is invariant under linear transformations of the parameter vector, whereas A- and E-optimality are variable.

The limitation to these and comparably simple cost functions were dictated by a lack of efficient optimization algorithms. All mentioned functions depend analytically on the entries of $\overline{M}$ and may be efficiently minimized using, e.g. Kiefer–Wolfowitz theory (Kiefer, 1974) [2], which was already available in the early 1970s.

In practical applications the number of input data is often sufficiently large, but the assumption of the model order tending to infinity is not very realistic. Therefore it is advisable to use a covariance formula which is asymptotic only in the number of data, but valid for finite model order. Such formulas were known for decades (Ljung, 1999) [3], but input design problems with cost functions involving them were too difficult to solve. This changed with the appearance of powerful convex optimization methods. In Hildebrand and Gevers (2003) [4], a more complicated input design criterion based on the more exact covariance formula was treated whose minimization was conducted by the ellipsoid method (Boyd, El Ghaoui, Feron, & Balakrishnan, 1994) [5]. This is a particularly robust method which requires only information about the gradient of the cost function and supporting hyperplanes to the feasible set of search parameters, and can hence minimize a large variety of cost functions. Sometimes input design problems or their relaxed versions can be solved by semidefinite programming (see e.g. Bombois, Scorletti, Gevers, Hildebrand, & van den Hof, 2004) [6]. A summary of such methods is given in Jansson (2004) [7].

This paper introduces adaptive parameter identification technique, and then the optimal input design problem is generally defined. The procedure of obtaining the approximate white noise is expressed, and finally, the further albinism approximate white noise to the shaking table experimental system is implemented.

II. ADAPTIVE PARAMETER IDENTIFICATION TECHNIQUE

The presented study considers the parameter identification of the following dynamic system:

$$\dot{x}_i = f_i(X,U) + \theta_i^T g_i(X,U), i = 1,\cdots,n. \quad (1)$$

978-1-4244-1676-9/08 /$25.00 ©2008 IEEE

where $X=[x_1,\cdots,x_n]\in R^n$ is the state vector; $U\in R^p$ is the excitation signal (or input); $\theta_i\in R^{n_i}$, $n_i\geq 0$, is the unknown parameter vector; and $f_i(\cdot)\in R$ and $g_i(\cdot)\in R^{n_i}$ are the known scalar and vector functions, respectively.

The present study presents a parameter estimator of the following form:

$$\dot{\hat{x}}_i=f_i(X,U)+\theta_i^T g_i(X,U)-\lambda_i e_i. \tag{2}$$

where $\theta_i\in R^{n_i}$ is the estimated parameter vector, λ_i is a positive number to be specified, and $e_i\equiv\hat{x}_i-x_i$ is the state error. The adaptive law for adjusting $\hat{\theta}_i$ (or ϕ_i) is given by

$$\dot{\hat{\theta}}_i=\dot{\phi}_i=-\gamma_i e_i g_i(X,U),\gamma_i>0. \tag{3}$$

The stability of the parameter estimator can be achieved by defining the parameter error ϕ_i as follows:

$$\phi_i=\hat{\theta}_i-\theta_i. \tag{4}$$

and then a quadratic Lyapunov function is chosen by

$$V=\frac{1}{2}\sum_{i=1}^{n}\left(e_i^2+\frac{\phi_i^T\phi_i}{\gamma_i}\right). \tag{5}$$

Differentiating the above function along the trajectory of (1)-(3) is given by

$$\dot{V}=\sum_{i=1}^{n}\left(-\lambda_i e_i^2\right)<0. \tag{6}$$

which implies that

$$\lim_{t\to\infty}e_i(t)=0,i=1,\cdots,n. \tag{7}$$

However, the asymptotic stability of the point (e_i,ϕ_i), cannot be concluded directly since $\dot{V}$ is only negative-semidefinite. In practice, the convergence of $\phi_i(t)$ to zero is dependent on the characteristic of the excitation signal g_i (excited by U), which is referred to as persistent excitation (PE), highly important to the issue of parameter identification [11–16]. According to one of the several equivalent conditions given in Theorem 2.16 in Ref. [12]: $\phi_i=0$ in Eq. (3) is uniformly asymptotically stable if, and only if, the piecewise-continuous uniformly bounded function g_i satisfies the inequality:

$$\int_t^{t+T_0}g_i(\tau)g_i^T(\tau)d\tau\geq\alpha_i I,\forall t\geq t_0. \tag{8}$$

for positive constants t_0, T_0, and α_i. However, the PE condition on g_i is not a verifiable condition since the signal g_i is generated inside the adaptive system and is unknown a priori; so there is no way for us to check whether or not it is persistently exciting. The literature [13] points out that the ultimate result of such investigation is that the input $U(t)$ should be "sufficiently rich" to guarantee that g_i is persistently exciting. For example, if $u(t)\in R$ is the sum of distinct sinusoids, then it should have at least n_i distinct frequencies. Therefore, the intuition behind the PE condition is that the input should be rich enough to excite all modes of the system [13].

III. THE OPTIMAL INPUT

In the system identification, for comparing with different experiment, an unification measurement of evaluating experiment "the quality", which should concern with the input - output signal fitting parameter precision. Rao and Cramer proved that the covariance of parameter estimation satisfies the inequality with some conditions:

$$\operatorname{cov}\hat{\theta}\geq M_\theta^{-1}. \tag{9}$$

Where, M_θ is called Fisher information matrix. When the inequality takes the equal sign, called the parameter estimation achieving the smallest variance bound (Rao_Cramer low bound), and the parameter estimation precision achieves the highest. During an experiment, one always takes

$$J=-\log(\det(M_\theta)). \tag{10}$$

or

$$J=tr(M_\theta^{-1}). \tag{11}$$

as the optimal design criterion.

Considers MA model:

$$y_t=b_1u_{t-1}+b_2u_{t-2}+\cdots+b_nu_{t-n}+\varepsilon_t. \tag{12}$$

and supposes the noise ε_t obedience normal distribution N (0, σ^2), with

$$\beta^T=(b_1,b_2,\cdots,b_n,\sigma^2)=(\theta^T,\sigma^2). \tag{13}$$

Its Fisher information matrix is [8]:

$$M_\beta=\begin{bmatrix}\frac{N}{\sigma^2}\cdot\Gamma & 0\\ 0 & \frac{N}{2\sigma^4}\end{bmatrix}=\begin{bmatrix}M_\theta & 0\\ 0 & \frac{N}{2\sigma^4}\end{bmatrix}. \tag{14}$$

Where

$$\Gamma = \frac{1}{N}\sum_{t=1}^{N}\begin{bmatrix} u_{t-1}^2 & u_{t-1}u_{t-2} & \cdots & u_{t-1}u_{t-n} \\ u_{t-2}u_{t-1} & u_{t-2}^2 & \cdots & u_{t-2}u_{t-n} \\ \vdots & \vdots & \vdots & \vdots \\ u_{t-n}u_{t-1} & u_{t-n}u_{t-2} & \cdots & u_{t-n}^2 \end{bmatrix}. \quad (15)$$

In which the input signal obey power constraint:

$$\max_i\left\{\frac{1}{N}\sum_{t=1}^{N}u_{t-i}^2\right\}=1, \qquad i=1,\cdots,n. \quad (16)$$

According to the known result, minimizing J is equivalent to seek the sequence $\{u_t, t=1-n,\cdots,N-1\}$, which causes

$$\frac{1}{N}\sum_{t=1}^{N}u_{t-i}u_{t-j}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}. \quad (17)$$

If input signal satisfies this request, then parameter estimation achieves minimum value of the smallest variance with effective algorithm:

$$\operatorname{cov}\hat{\theta}=M_{\theta}^{-1}=\frac{\sigma^2}{N}\cdot I. \quad (18)$$

When N be large enough, the input signal can be replaced approximately by white noise sequence.

IV. CREATION OF WHITE NOISE

As stated previously, once a mathematical model and a suitable parameter identification technique have been determined, the result of the parameter identification process is contingent upon the particular choice of excitation signal. In order to guarantee the convergence of the estimated parameters, it may be necessary to ensure that the excitation signal satisfies the PE conditions. In practical applications, it is more useful to analyze the characteristic of the excitation signals in their frequency domain. This method is commonly employed in practice since decomposing a signal into a set of harmonic signals is a well-established principle among scientists and engineers. In theory, the optimal excitation signal will be uncorrelated, or white, i.e. its spectrum will be broadband over all frequencies. Therefore, the output contains maximum information about the dynamic modes of the system. Ideally, the optimal excitation signal would be in the form of an impulse signal, which has a flat power spectrum. However, generating this signal is difficult, and hence engineers generally specify random sequences, PRBS, or swept-frequency sinusoidal signals as excitation signals. Especially for shaking table experimental system, white noise is the most fitting input excitation signal for this system.

The white noise process is a simple stochastic process. Strictly saying, it is a constant stationary random process with average value zero and the spectral density non-zero, or it is an idealized stochastic process which is composed by a series of non-correlated random variables. The white noise process does not have "the memory", in other words, the *t*-time value had nothing to do with the past value at *t*-time beforehand, and will not affect future value of the *t*-time later.

Defining White noise process: If average value of the stochastic process $w(t)$ is zero, the autocorrelation function is

$$R_w(t)=\sigma^2\delta(t). \quad (19)$$

Where $\delta(t)$ is Dirac function, namely

$$\delta(t)=\begin{cases}\infty & t=0\\ 0 & t\neq 0\end{cases}. \quad (20)$$

and

$$\int_{-\infty}^{\infty}\delta(t)dt=1. \quad (21)$$

Thus this stochastic process is called as white noise process.

White noise which conforms to the above definition strictly is only a theoretically abstract, and is impossible to implement in practical applications. The stochastic process can be regarded as a white noise approximately when its average power is in uniformly distribution over the useful frequency band.

According to the white noise definition, the Fourier peak value spectrum can be obtained with the power spectrum root, and transform the Fourier transformation the real and the imaginary with defining willfully stochastic phase spectrum and the Fourier peak value spectrum, then makes the Fourier inverse transformation, the approximate white noise can be finally obtained. The generated white noise expression is given by

$$A(t)=FFT^{-1}[\sqrt{S(\omega)}e^{i\phi(\omega)}]. \quad (22)$$

$$S(\omega)=\begin{cases}1 & \omega_0\le\omega\le\omega_1\\ 0 & others\end{cases}. \quad (23)$$

In which: A(t) is the approximate white noise signal; $S(\omega)$ is the power spectrum of the frequency range $[\omega_0,\omega_1]$; $\phi(\omega)$ is the stochastic phase spectrum, random numbers between $0\sim 2\pi$; i is the imaginary number, namely $\sqrt{-1}$.

As shown in Figure 1, x(n) is white noise sequence with the length 100 second, the sampling frequency 400Hz which must be above 225Hz,and the frequency range $0.5\sim 90Hz$; psd(f) expressed the power spectral density function, R(k) expressed autocorrelation function of the sequence x(n), or auto-covariance function. This power spectral density function undulates along with the frequency change from figure 1, it corresponds to the R(k) which undulates with encompassment time axis, this explained the white noise "white" degree.

Fig. 2 and Fig. 3 contrast the power spectral density and the autocorrelation function of the white noise with different data

length. It can be observed that the undulation of power spectral density is smaller, and the autocorrelation function is closer to $\delta(t)$ function with longer data length.

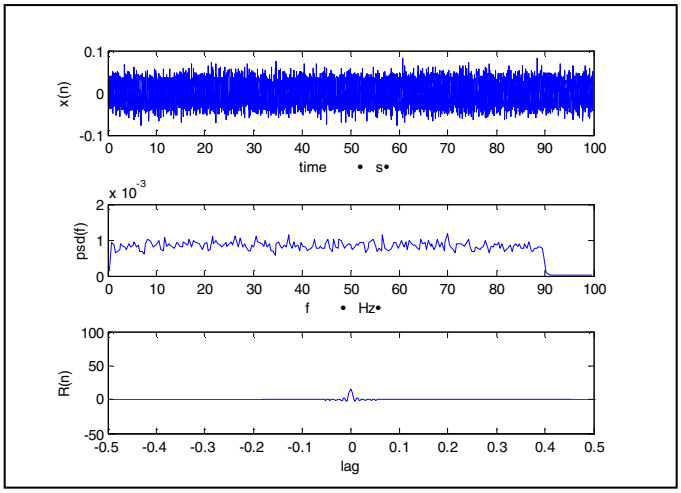

Fig.1 the approximate white noise sequence with time interval, power spectral density and autocorrelation function

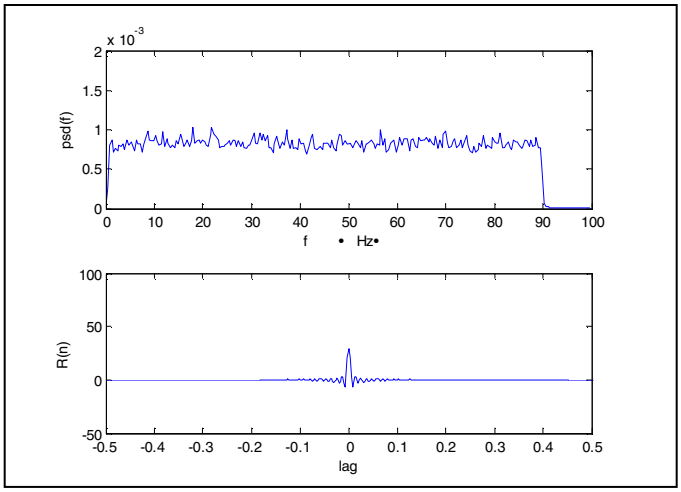

Fig.2 power spectral densities and the autocorrelation function with the length 200 second

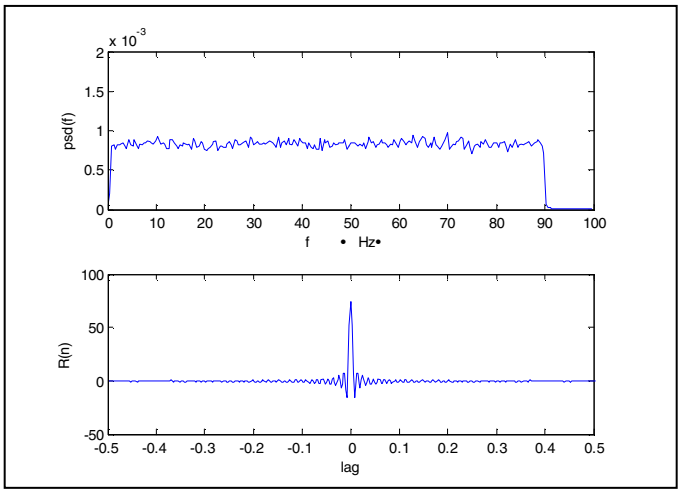

Fig.3 power spectral densities and the autocorrelation function with the length 500 second

V. APPROXIMATE WHITE NOISE FURTHER ALBINISM

In shaking table experimental system, experimental time is impossible infinite with certain limitation, so the input signal also has time limitation. Along with increase in experimental length N, parameter estimation precision goes up, and error of estimation variance drops, therefore N should be chosen as longer as possible. But as a result of the actual condition limits, the probability with drifts and disturbance increases with longer N, and the computation load also enlarges, therefore the N is not suitably taken too longer. Thus, the choice of N should consider the specific identification situation, and follow a cardinal principle that experiment time should be 10 times greater than the identified system main time-constant at least, namely

$$NT_0 \geq 10T_a . \tag{24}$$

Where, T_0 is the sample period, T_a is the main time-constant of the identified system.

How to cause the approximate white noise to be closer to the ideal white noise with the same data length is to be solved in the next section.

One may see by literature [9-10] that the best method white processing to the sequence is the double stochastic interchange optimal algorithm which is first used by Hunter and Kearney. The basic principle of this algorithm is to seek N-length data most superior arrangement of sequence x(n) with stochastic search optimization method, and makes autocorrelation function r(k) of rearrangement sequence to approach $\delta(t)$ function. This algorithm looping execution step is as follows:

1st step: Draws 2 data-points of random sequence $x_{i-1}(n)(i=1,2,\cdots)$ and exchanges their position mutually, results in sequence $x_i(n)$ (this is meaning of double stochastic interchange);

2nd step: Computes autocorrelation function of sequence $x_i(n)$

$$r_i(k) = \frac{1}{N}\sum_{n=0}^{N-k-1} x_i(n)x_i(n+k) \qquad k = 0,1,\cdots,N-1. \tag{25}$$

($x_0(n) = x(n)$ may be used to calculate the initial autocorrelation function $r_0(k)$);

3rd step: Calculates the sum of squares

$$S_i = \sum_{k=1}^{N-1}[r_i(k)]^2 \qquad i = 0,1,2,\cdots \quad (S_0 \text{ is initial value}). \tag{26}$$

4th step: If $S_i < \varepsilon$, or $i = n_i$, then stop;

5th step: If $S_i < S_{i-1}$, then returns 1st step; Otherwise, gives up the i-th stochastic double interchange in the 1st step, and returns the 1st step.

Where, ε is the predetermined admissible error, N_i is the predetermined biggest cycle-index of execution operation.

As Figure 3 shown, the white noise sequence is of the length 100 second, the sampling frequency 400Hz, and the frequency range 0.5~90Hz through albinism result of above algorithm. One may see that the undulation of white noise power spectral density function is smaller after the albinism process, and the autocorrelation function approaches $\delta(t)$ function. This can solve the question of the shaking table experimental system with the input signal length limitation.

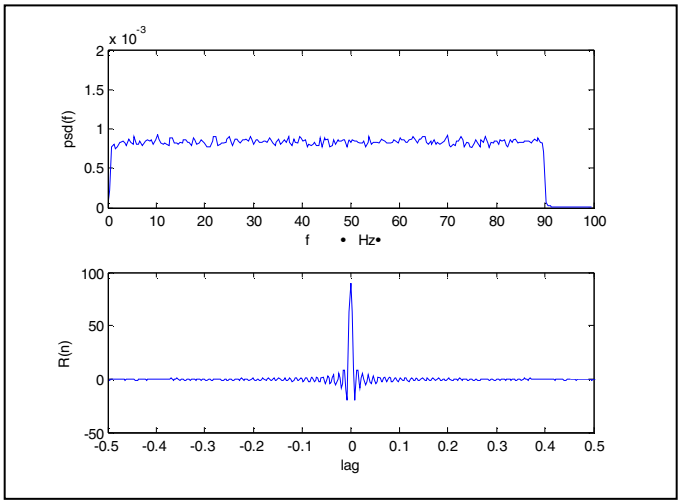

Fig.4 power spectral density and autocorrelation function with albinism processing

VI. CONCLUSION

This paper presents a good idea for the question of the shaking table experimental system with the input signal length limitation, and has made the input driving signal most greatly possibly satisfy the PE condition in the data length limitation. This can improve the doubly stochastic interchange optimized algorithm, by which the operation speed is expedited during a shaking table test. However, for the experimental system, it is worth studying how long the driving time need in test in the future.

ACKNOWLEDGMENT

This work is partially supported by National Nature Science Foundation of China (Grant No. 60574052), National Nature Science Foundation of China (Grant No. 50578046), Technical Department Special Project of China (Grant No. 2004CCA03300), Guangzhou Science and Technology Plan Project (Grant No. 2004Z1-E0051).

REFERENCES

[1] Mehra, R. Optimal input signals for parameter estimation in dynamic systems—survey and new results. IEEE Transactions on Automatic Control, AC-19(6), pp. 753–768, 1974.

[2] Kiefer, J. General equivalence theory for optimum designs (approximate theory). Annals of Statistics, 2(5), pp. 849–879, 1974

[3] Ljung, L. (1999). System identification: Theory for the user (2nd ed.), System sciences series. Englewood Cliffs, NJ: Prentice-Hall.

[4] Hildebrand, R., & Gevers, M. Identification for control: Optimal input design with respect to a worst-case -gap cost function. SIAM Journal of Control Optimization, 41(5), pp. 1586–1608, 2003

[5] Boyd, S., El Ghaoui, L., Feron, E., & Balakrishnan, V. Linear matrix inequalities in system and control theory. Philadelphia: SIAM. 1994.

[6] Bombois, X., Scorletti, G., Gevers, M., Hildebrand, R., & van den Hof, P. Cheapest open-loop identification for control. In Proceedings of the 43rd IEEE CDC, Atlantis, Bahamas, USA, pp. 382–387, 2004.

[7] Jansson, H. Experiment design with applications in identification for control. Ph.D. Thesis, Royal Institute of Technology, Stockholm, Sweden. 2004.

[8] Guangwen Han. Identification and parameter estimation. National defence industry press, 1980.

[9] Hunter I W, Keamey R E. Generation of Stochastic Sequences for System Identification [C]. Proc. 35th Ann Conf. Eng. Med. Biol, Philadelphia, PA, September, (24) pp. 72, 1982.

[10] Hunter I W, Kearney R E. Generation of Random Sequences with Jointly Specified Probability Density and Autocorrelation Functions [J]. Biol. Cybern,, (47) pp. 141-146, 1983.

[11] J.S-C. Yuan, W.M. Wonham, Probing signals for model reference identification, IEEE Transactions on Automatic Control 22 pp. 530-538, 1977

[12] K.S. Narendra, A.M. Annaswamy, Stable Adaptive Systems, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[13] H.K. Khalil, Nonlinear System, 2nd edn, Prentice-Hall, Englewood Cliffs, New Jersey, 1996.

[14] R.R. Bitmead, Persistence of excitation conditions and the convergence of adaptive schemes, IEEE Transactions on Information Theory 30 pp. 183–191, 1984.

[15] S. Boyd, S. Sastry, Necessary and sufficient conditions for parameter convergence in adaptive control, Automatica 22 pp. 629–639, 1986.

[16] K.S. Narendra, A.M. Annaswamy, Persistent excitation of adaptive systems, International Journal of Control 45 pp. 127–160, 1987.