IDENTIFICATION OF MULTI-DIMENSIONAL SYSTEM BASED ON A NOVEL CRITERION

Yue Zhao, Kueiming Lo
School of Software, Tsinghua University, Key Lab for ISS, MOE China, Tsinghua University, Beijing 100084, P. R. China
yue-zhao07@mails.tsinghua.edu.cn, gluo@tsinghua.edu.cn

Wook-Hyun Kwon
School of Electrical Engineering, Seoul National University, Seoul 151-742, Korea
whkwon@cisl.snu.ac.kr

Keywords: Multi-dimensional system, identification criterion, ARMAX model, recursive algorithm.

Abstract: Most system recursive identification algorithms are based on the prediction error (PE) criterion. Such a recursive algorithm only considers the present estimation residual error instead of all estimation residuals. It would result in large estimation error when the signal noise disturbs strongly. In this paper, a new identification criterion is proposed. It considers both the errors between the actual outputs and the estimation result and the difference of each estimation error. Under this criterion, a new recursive algorithm MSDCN (Multi-dimensional System Disturbed by Color Noise) is proposed. For multi-dimensional systems, weighting different values on the estimation errors and the difference of each error, MSDCN could both decrease the estimation errors and got smooth prediction curves. Several simulation examples are given to illustrate the method's anti-disturbance performance.

1 INTRODUCTION

There already have many recursive algorithms for modeling systems disturbed by white noises (Andersson and Broman 1998, Ljung 1999, Griffith Jr 1999, Mershed 2000, etc.). The most characteristic feature of disturbance, however, is that its value is not known beforehand. A physical system is affected by many factors such as color noise disturbance and an un-modeled structure. These types of algorithms encounter difficulties when the measurements are disturbed (Kuo 2000, Trump 2001). For multi-discrete systems disturbed by color noise, current identification algorithms cannot give precise estimate results either (Schoukens 1991, Ljung 1985).

Efficiency modeling depends on the choice of identification criterion. Combining frequency-domain method and time-domain method the GPE criterion was proposed for interference systems (Lo and Kwon 2002, 2003). Some extended recursive algorithms (ERA) were put forward based on the GPE criterion. (Lo and Kimura 2003, Lo *et al.* 2006, and Lo and Huang 2006) Usually, a GPE criterion contains a weighting matrix, in which there are many parameters for free choice. When the weighting matrix is the identity, the ERA becomes the recursive least squares algorithm (Lo and Kimura 2003). In addition, the standard Yule-Walker method uses identity matrix as weighting matrix. It may have poor accuracy, and increasing the dimension of the weighting matrix may well degrade the accuracy (Stoica, Friedlander and Soderstrom 1987). Stoica and Jansson (2001) proposed another method which derived the optimal weight in a simple way and guaranteed the optimal weighting matrix to be consistent and non-negative definite, while still the choice of weighing matrix is hard. Furthermore, for all the above raised implementations, a positive definite weighting matrix must be weighted out in order to get a reliable estimate. The optimal weight in general depends on unknown quantities and hence must be itself estimated before its use become possible.

Considering both the prediction errors and the difference of each error, performance criterion in this paper applies different weights. One part is the errors between the actual outputs and the other is the estimation result and the difference of each estimation error. Further, this paper develops a recursive weighting matrix for fast calculation in recursive algorithms. In this matrix, the values of each element depends on the weights of the performance function. Based on this new performance function, a new recursive algorithm for multi-system identification, MSDCN, is proposed.

MSDCN algorithm is new not only because it is based on a new performance function, but also because it is based on a new concept of estimating the noise part separately. Its two-step estimation feature also make it a novel method. The MSDCN algorithm has good anti-disturbance. For multi-dimensional system, using the novel criterion to do the complex estimation in each dimension of parameters, MSDCN can give more precise results than current algorithms do.

In the second part of this paper, the extended recursive algorithm for multi-dimensional systems is introduced; The third part proposes a novel identification criterion; In the fourth part a recursive algorithm for multi-dimensional system, MSDCN, is proposed; An analysis of the performance of new criterion and new algorithm is given in the fifth section, and then its simulation results are compared with other algorithms.

2 EXTENDED RECURSIVE ALGORITHM FOR MULTI-DIMENSIONAL SYSTEM

Consider system:

$$A(q)y(t) = B(q)u(t) + w(t) \tag{1}$$

where, $y(t)$ and $w(t)$ are p-dimensional vectors, $u(t)$ is m-dimensional vector. $y(t)$, $u(t)$, and $w(t)$ are system output, input, and noise respectively. $A(q)$, $B(q)$ are the backward operator q^{-1} polynomial expression

$$A(q) = I + \sum_{k=1}^{n_a} A_k q^{-k}, \quad B(q) = \sum_{k=1}^{n_b} B_k q^{-k}.$$

Denote:

$$\begin{aligned} \theta &= (A_1, A_2, \cdots, A_{n_a}, B_1, B_2, \cdots, B_{n_b})^T \\ \varphi_t &= (-y^T(t-1), \cdots, -y^T(t-n_a), \\ &\quad u^T(t-1), \cdots, u^T(t-n_b))^T \end{aligned}$$

where $\theta \in R^{(n_a \cdot p + n_b \cdot m) \times p}$ is a matrix formed by the system part parameters. $\varphi_t \in R^{n_a \cdot p + n_b \cdot m}$ is a regression vector. Then (1) is rewritten as

$$y(t) = \theta^T \varphi_t + w(t) \tag{2}$$

Denote ε_t be the prediction error of system output y(t):

$$\varepsilon_t = y(t) - \theta^T \varphi_t, \tag{3}$$

Then the identification criterion can be expressed as:

$$J(N) = tr([\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N]^T Q(N) [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N]) \tag{4}$$

where trA represents the trace of matrix A. The weighting matrix $Q(N) \in \mathbf{R}^{N \times N}$ is a symmetrical positive definite matrix and expressed as:

$$Q(N) = \begin{pmatrix} Q(N-1) & \alpha(N) \\ \alpha(N)^T & q_N \end{pmatrix}, \quad t = 1, 2, \cdots, N$$

where $\alpha(N) \in \mathbf{R}^{N-1}$. Denote:

$$\Phi(N) = (\varphi_1, \varphi_2, \cdots, \varphi_N)^T Y(N) = (y_1, y_2, \cdots, y_N)^T$$

Then the identification criterion (4) can be expressed as:

$$J(N) = tr([Y(N) - \Phi(N)\theta]^T Q(N) [Y(N) - \Phi(N)\theta] \tag{5}$$

Since

$$\begin{aligned} J(N) &= tr(Y^T(N)Q(N)Y(N)) - tr(\theta^T \Phi^T(N)Q(N)Y(N)) \\ &\quad - tr(Y^T(N)Q(N)\Phi(N)\theta) + tr(\theta^T \Phi^T(N)Q(N)\Phi(N)\theta) \end{aligned}$$

The gradient of J(N) is:

$$\begin{aligned} \frac{\partial J(N)}{\partial \theta} &= \frac{\partial}{\partial \theta}[tr(Y^T(N)Q(N)Y(N)) \\ &- tr(\theta^T \Phi^T(N)Q(N)Y(N)) \\ &- tr(Y^T(N)Q(N)\Phi(N)\theta) \\ &+ tr(\theta^T \Phi^T(N)Q(N)\Phi(N)\theta(N))] \\ &= \frac{\partial}{\partial \theta}[tr(Y^T(N)Q(N)Y(N))] - \\ &\frac{\partial}{\partial \theta}[tr(\theta^T \Phi^T(N)Q(N)Y(N))] \\ &- \frac{\partial}{\partial \theta}[tr(Y^T(N)Q(N)\Phi(N)\theta)] \\ &+ \frac{\partial}{\partial \theta}[tr(\theta^T \Phi^T(N)Q(N)\Phi(N)\theta(N))] \\ &= -(Y^T(N)Q(N)\Phi(N))^T - \Phi^T(N)Q(N)Y(N) + \\ &(\Phi^T(N)Q(N)\Phi(N) + \Phi^T(N)Q(N)^T\Phi(N))\theta(N) \end{aligned}$$

and

$$\frac{\partial J^2(N)}{\partial \theta^2} = \Phi^T(N)Q(N)^T\Phi(N)$$

where $Q(N) = Q^T(N)$ and $\frac{\partial J^2(N)}{\partial \theta^2}$ is positive and definite. Let

$$\frac{\partial J(N)}{\partial \theta} = 0$$

which minimizes $J(N)$ and yields:

$$\theta(N) = [\Phi(N)^T Q(N)\Phi(N)]^{-1}\Phi(N)^T Q(N)Y(N) \tag{6}$$

$t=1,2,\cdots$ Then at time t, denote:

$$\begin{aligned} P_t &= \Phi_t^T Q_t \Phi_t \\ a_t &= 1+\varphi_t^T P_{t-1}^{-1}\Phi_{t-1}^T \alpha_t \\ \sigma_t &= q_t - \alpha_t^T \Phi_{t-1} P_{t-1}^{-1} \Phi_{t-1}^T \alpha_t \\ b_t &= a_t + a_t^{-1}\sigma_t \varphi_t^T P_{t-1}^{-1}\varphi_t \end{aligned}$$

Then we can get:

Theorem 1

$$\begin{cases} P_t^{-1} = & P_{t-1}^{-1} \\ & - b_t P_{t-1}^{-1}(\varphi_t \beta_t^T \Phi_{t-1} + \Phi_{t-1}^T \beta_t \varphi_t^T) P_{t-1}^{-1} \\ & + \frac{1}{a_t b_t} P_{t-1}^{-1}(\varphi_t^T P_{t-1}^{-1} \varphi_t \Phi_{t-1}^T \beta_t \beta_t^T \Phi_{t-1} \\ & - \sigma_t \varphi_t \varphi_t^T) P_{t-1}^{-1} \\ \theta_t = & \theta_{t-1} \\ & + \frac{1}{a_t b_t} P_t^{-1}(\beta_t \Phi_t^T \beta_t + \sigma_t \varphi_t)(y_t - \theta_{t-1}^T \varphi_t) \\ & + \frac{1}{a_t b_t} P_{t-1}^{-1}(\beta_t \varphi_t - \varphi_t^T P_{t-1}^{-1} \varphi_{t-1}^T \beta_t)\beta_t^T (Y_{t-1} \\ & - \Phi_{t-1}\theta_{t-1}) \end{cases}$$

As Extended Recursive Algorithm for multi-dimensional system.

3 IDENTIFICATION CRITERION

New identification criterion is:

$$J(N)=\lambda \sum_{t=1}^{N} \parallel \varepsilon_t \parallel^2 + \mu \sum_{t=1}^{N-1} \parallel \varepsilon_{t+1}-\varepsilon_t \parallel^2 \quad (7)$$

where ε_t represents the estimation errors at time t, in multi-dimensional system, ε_t is p-dimensional vector, which dimension is the same as the output y(t).

λ represents the weight on the estimation errors

μ represents the weight on the difference between each estimation error.

Transforming equation (7), we can get

$$\begin{aligned} J(N) &= \sum_{t=1}^{N}(\lambda \varepsilon_t^T \varepsilon_t) + \sum_{t=1}^{N-1}\mu(\varepsilon_{t+1}-\varepsilon_t)^T(\varepsilon_{t+1}-\varepsilon_t)) \\ &= \sum_{t=1}^{N} tr(\lambda \varepsilon_t \varepsilon_t^T) + \sum_{t=1}^{N-1} \mu tr(\varepsilon_{t+1}-\varepsilon_t)(\varepsilon_{t+1}-\varepsilon_t)^T \\ &= tr\left((\lambda+2\mu)\sum_{t=1}^{N}\varepsilon_t\varepsilon_t^T - \mu\sum_{t=1}^{N-1}\varepsilon_{t+1}\varepsilon_t^T - \mu \sum_{t=1}^{N-1}\varepsilon_t \varepsilon_{t+1}^T\right) \end{aligned}$$

Assume $E(N)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_N)^T \in N\times p$ and

$$Q(N)=\begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1N} \\ q_{21} & q_{22} & \cdots & q_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ q_{t1} & q_{t2} & \cdots & q_{NN} \end{pmatrix} \in N \times N,$$

in which,

$$\begin{cases} q_{i,i} = \lambda + 2\mu \\ q_{i+1,i} = -\mu & i=1,2,\cdots,N. \\ q_{i,i+1} = -\mu \end{cases}$$

Then equivalent with equation (3), we can get

$$J(N)=trE(N)^T Q(N) E(N) \quad (8)$$

as the new expression for the performance criterion.

Thus at time t, the weighting matrix Q_t can be expressed in this recursive form:

$$Q_t = \begin{pmatrix} Q_{t-1} & \beta_t \\ \beta_t^T & p_t \end{pmatrix}, \qquad t=1,2,\cdots,N$$

in which, $\beta_t=[0,0,\cdots,-\mu]\in \mathbf{R}^{t-1}$ and $p_t=\lambda+2\mu$.

4 RECURSIVE ALGORITHM FOR MULTI-DIMENSIONAL SYSTEM

Consider the ARMAX model

$$A(q)y(t)=B(q)u(t)+C(q)w(t) \quad (9)$$

where, $y(t)$ and $w(t)$ are p-dimensional vectors, $u(t)$ is m-dimensional vector. $y(t)$, $u(t)$, and $w(t)$ are system output, input, and color noise respectively. $A(q)$, $B(q)$, $C(q)$is the backward operator q^{-1} polynomial expression

$$A(q)=I+\sum_{k=1}^{n_a} A_k q^{-k}, \quad B(q)=\sum_{k=1}^{n_b} B_k q^{-k},$$
$$C(q)=I+\sum_{k=1}^{n_c} C_k q^{-k}.$$

where $A_k, C_k \in p\times p$, and $B_k \in p\times m$. Denote:

$$\theta=(A_1,A_2,\cdots,A_{n_a},B_1,B_2,\cdots,B_{n_b})^T,$$

$$\varphi_t=(-y^T(t-1),\cdots,-y^T(t-n_a),u^T(t-1),\cdots,u^T(t-n_b))^T,$$

$$\rho=(C_1,C_2,\cdots,C_{n_c})^T,$$
$$\phi_t=(w(t-1),w(t-2),\cdots,w(t-n_c))^T.$$

in which $\theta \in R^{[n_a\cdot p+n_b\cdot m]\times p}$ is the matrix formed by the system part parameters; and $\rho \in R^{[n_c\cdot p]\times p}$ is the matrix formed by the system's noise part parameters.

Here we introduce a two-step algorithm to do the estimation for system (9).

Step 1: Transform (9) into:

$$y(t)-\rho_{t-1}^T\phi_t=\theta_t^T\varphi_t+w(t) \quad (10)$$

Estimate the system parameters θfor system (10). For the identification of the parameters of system θ, use the following measurements:

$$J(t) = tr(\lambda \sum_{t=1}^{N} \parallel \varepsilon_t \parallel^2 + \mu \sum_{t=1}^{N-1} \parallel \varepsilon_{t+1} - \varepsilon_t \parallel^2). \quad (11)$$

Step 2: Transform (10) into:

$$y(t) - \theta_t^T \varphi_t = \rho_t^T \phi_t + w(t). \quad (12)$$

Estimate the parameter ρ of filter $C(q)$ of system (9).

Denote

$$\Phi_t = (\varphi_1, \varphi_2, \cdots, \varphi_t)^T, Y_t = (y_1, y_2, \cdots, y_t)^T,$$
$$g_t = y_t - \rho_{t-1}^T \phi_t, G_t = (g_1, g_2, \cdots, g_t)^T$$

And according to equation (8), the novel criterion $J(t)$ can be expressed as:

$$\begin{aligned} J(t) &= trE(t)^T Q_t E(t) \\ &= tr[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_t]^T Q_t [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_t] \\ &= tr[G_t - \Phi_t \theta]^T Q_t [G_t - \Phi_t \theta] \end{aligned}$$

If $\Phi_t^T Q_t \Phi_t$ is not singularthen minimize J_t to get the system's parameter vector θ_t's optimal solution:

$$\theta_t = [\Phi_t^T Q_t \Phi_t]^{-1} \Phi_t^T Q_t G_t, \quad t = 1, 2, \cdots, N. \quad (13)$$

Similar with Extended Recursive Algorithm, we can get MSDCN algorithm for ARMAX model (9):

$$\left\{ \begin{aligned} \theta_M(t) = \ & \theta_M(t-1) \\ & + \frac{1}{a_t b_t} P_t^{-1} (\beta_t \Phi_t^T \beta_t + \sigma_t \varphi_t)(y_t \\ & - \rho_M(t-1)^T \phi_t - \theta_M(t-1)^T \varphi_t) \\ & + \frac{1}{a_t b_t} P_{t-1}^{-1} (\beta_t \varphi_t - \varphi_t^T P_{t-1}^{-1} \varphi_{t-1}^T \beta_t) \beta_t^T (G_{t-1} \\ & - \Phi_{t-1} \theta_M(t-1)) \\ P_t^{-1} = \ & P_{t-1}^{-1} \\ & - b_t P_{t-1}^{-1} (\varphi_t \beta_t^T \Phi_{t-1} + \Phi_{t-1}^T \beta_t \varphi_t^T) P_{t-1}^{-1} \\ & + \frac{1}{a_t b_t} P_{t-1}^{-1} (\varphi_t^T P_{t-1}^{-1} \varphi_t \Phi_{t-1}^T \beta_t \beta_t^T \Phi_{t-1} \\ & - \sigma_t \varphi_t \varphi_t^T) P_{t-1}^{-1} \\ \rho_M(t) = \ & \rho_M(t-1) + \frac{R_{t-1}^{-1} \phi_t}{1 + \phi_t^T R_{t-1} \phi_t} (y_t \\ & - \theta_M(t)^T \varphi_t - \rho_M(t-1)^T \phi_t) \\ R_t = \ & R_{t-1} + \phi_t \phi_t^T \\ \widehat{w}(t) = \ & y(t) - \rho_M(t)^T \phi_t - \theta_M(t)^T \varphi_t \end{aligned} \right. \quad (14)$$

in which,

$$\left\{ \begin{aligned} & p_t = \lambda + 2\mu \\ & \beta_t = [0, 0, \cdots, -\mu] \in \mathbf{R}^{t-1} \end{aligned} \right.$$

Initially, θ and ρ can be zero matrix; R_0 , P_0 together constitute the identity matrix.

5 SIMULATIONS

All the simulations were conducted in the same computation environment. The main criterions of the computer were: CPU 1.66GHz, RAM 1G bytes and with Windows XP OS.

Experiment. A Black-box model is as follows:

$$y(t) = \frac{B_1 q^{-1}}{I + A_1 q^{-1}} u(t) + \frac{I + C_1 q^{-1}}{I + A_1 q^{-1}} w(t)$$

The real parameters of the system were:

$$A_1 = \begin{pmatrix} 0.325 & -1 \\ 0.5 & -1.1 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} 1.9 & 3.7 \\ -0.4 & -0.8 \end{pmatrix},$$

$$C_1 = \begin{pmatrix} -0.7 & -0.1 \\ 0.8 & 0.9 \end{pmatrix}.$$

The color noise sequence $\{w(t)\}_1^N$ was a two dimensional vector, of which each dimension was composed of different sawtooth wave and random number generator with variance 2 and the input signal $u(t) = [1\ 2]$, each dimension of which was generated by a square generator. The experiment was conducted using the LS method and the algorithm MSDCN method with the sample number $N = 1000$. The results are shown in Figures 1,2,3,4,5,6 and the statistics are as follows:

(1)Least-Squares method

The system statistics results are:

$$\theta = \begin{pmatrix} 0.2889 & 0.0361 \\ -0.9823 & -0.0177 \\ 0.3855 & 0.1145 \\ -1.2366 & 0.1366 \\ 1.7990 & 0.1010 \\ 3.5981 & 0.1019 \\ -0.1692 & -0.2308 \\ -0.3384 & -0.4616 \\ -0.7499 & 0.0499 \\ 0.0907 & -0.1907 \\ 0.9270 & -0.1270 \\ 0.8506 & 0.0494 \end{pmatrix}.$$

The simulation results are shown in Figure 1, 2 and 3.

(2)MSDCN method

if we choose $\lambda = 0.8 and \mu = 0.2$,

then according to (7), $\beta_t = [0, \cdots, -0.2, -0.2]$ and $p_t = 1.2$.

Then the system statistics results are:

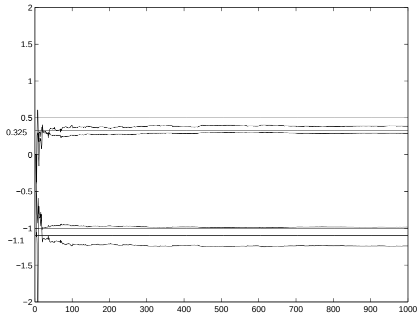

Figure 1: Estimation results of parameter A under ELS.

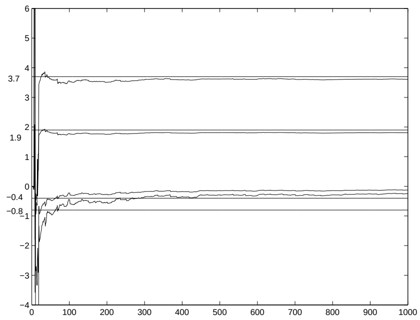

Figure 2: Estimation results of parameter B under ELS.

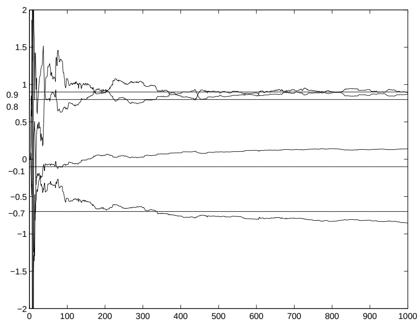

Figure 3: Estimation results of parameter C under ELS.

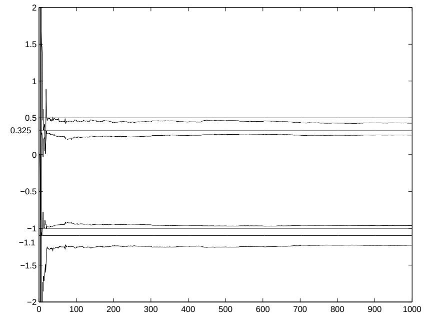

Figure 4: Estimation results of parameter A under MSDCN.

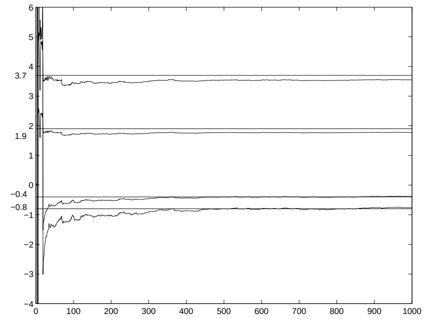

Figure 5: Estimation results of parameter B under MSDCN.

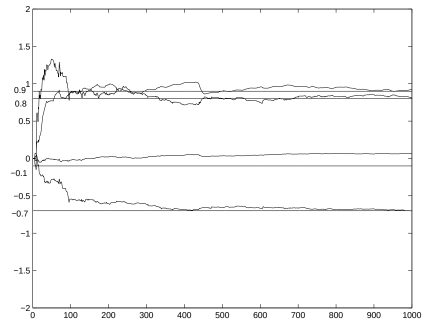

Figure 6: Estimation results of parameter C under MSDCN.

$$\theta = \begin{pmatrix} 0.2620 & 0.0630 \\ -0.9602 & -0.0398 \\ 0.4448 & 0.0552 \\ -1.2425 & 0.1425 \\ 1.7595 & 0.1405 \\ 3.5189 & 0.1811 \\ -0.4267 & 0.0267 \\ -0.8534 & 0.0534 \\ -0.6499 & -0.0501 \\ 0.0409 & -0.1409 \\ 0.8232 & -0.0232 \\ 0.9362 & -0.0362 \end{pmatrix}$$

The simulation results are shown in Figure 4, 5 and 6.

Through a comparison of the estimation results of ELS and MSDCN, we can see that the new performance function has proved to be efficient for the multi-dimensional systems. For the example given above, we weight 0.8 on the estimation errors and 0.2 on the difference between each estimation error, which means we want the estimation curves close to the actual values more than make the curves smooth. The weighting matrix is definite after β_t and p_t are fixed by λ and μ. Weighting different on λ and μ will result in different estimation results. Those figures show that MSDCN has both decreased the estimation errors and got smooth prediction curves.

6 CONCLUSIONS

This paper proposes a new identification criterion for multi-dimensional system disturbed by color noise, and further develops a recursive algorithm, MSDCN, based on it. Weighting both on the prediction errors and the difference of each prediction error, the identification criterion makes the weighting matrix definite in calculation. Based on this performance criterion, MSDCN is developed using a two-step method to estimate both the system parameters and the noise part. The MSDCN algorithm has high anti-disturbance performance in the prediction of multi-dimensional systems disturbed by color noise. It both decreased the estimation errors and got smooth prediction curves. The performance of the MSDCN algorithm was demonstrated by simulations.

ACKNOWLEDGEMENTS

This work is supported by the Funds NSFC60672110, NSFC60474026,and the JSPS Foundation.

REFERENCES

Andersson, A. and Broman, H. (1998), A second-order recursive algorithm with applications to adaptive filtering and subspace tracking. *IEEE Trans. Signal Processing*, vol. 46, pp. 1720C1725, June.

Griffith Jr, D. W. and Arce, G. R. (1999), A partially decoupled RLS algorithm for volterra filters. *IEEE Trans. Signal Processing*, vol. 47, pp. 579C582, Feb.

Kuo, C. J. , Jr, R. D. , Lin, C. Y. and Tsai, Y. C. (2000), Set-theoretic estimation based on a priori knowledge of the noise distribution. *IEEE Trans. Signal Processing*, vol. 48, pp. 2150C2156, July.

Lo, K. M. , *et al*. (2006), Empirical frequency-domain optimal parameter estimate for Black-box processes. *IEEE Transactions on Circuits and Systems-I: Regular Papers*, vol.53, No.2, 419-430.

Lo, K. M. , Kwon, W. H. (2003), New identification approaches for disturbed models. *Automatica*.vol.39, No.9, 1627-1634.

Lo, K. M. , Kimura, H. (2003), Recursive Estimation Methods for Discrete Systems. *IEEE Trans. On Circuits and Systems*.vol.49, NO.6, No.6, 439-446.

Ljung, L. (1985), On the estimation of transfer function. *Automatica*vol. 21, 677-696.

Ljung, L. (1999), System Identification: Theory for the User. *Upper Saddle River,NJ: Prentice-Hall*, 1999.

Mershed, R. and Sayed, A. J. (2000), Order-recursive RLS Laguerre adaptive filtering. *IEEE Trans. Signal Processing*, vol. 48, pp. 3000C3010, July.

Stoica, P., Friedlander, B.and Soderstrom, T.(1987) Optimal Intrumental Variable Multistep Algorithms for Estimation of the AR Parameters of an ARMA Process, *Int. J. Control* 45(1987), 2083-2107.

Stoica, P. and Jansson, M.(2001) Estimating Optimal Weights for Instrumental Variable Methods. *Digital Signal Processing - A Review Journal* vol. 11, no. 3, pp. 252-268, Jul. 2001.

Trump, T. (2001), Maximum likelihood trend estimation in exponential noise. *IEEE Trans. Signal Processing*, vol. 49, pp. 2087C2095, Sept.