Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Independent component analysis for nonminimum phase systems using H_∞ filters

Shuichi Fukunaga *,** **Kenji Fujimoto** *

* *Department of Mechanical Science and Engineering, Graduate Shool of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan*
** *Department of Monodukuri Engineering, Tokyo Metropolitan College of Industrial Technology, 1-10-40 Higashioi, Shinagawa-ku, Tokyo, 140-0011, Japan*

Abstract: This paper proposes an independent component analysis (ICA) method using H_∞ filters for nonminimum phase systems. In the basic ICA approach, the input signal is recovered by estimating the parameter of the inverse of the mixing system. If the system is nonminimum phase, the estimated parameter diverges due to the instability of the inverse. For this problem, a inverse filter is constructed based on an H_∞ filter in order to estimate the state of the given plant. The learning algorithm to estimate the parameter of the system is derived by minimizing the Kullback-Leibler divergence. Furthermore, a numerical simulation demonstrates the effectiveness of the proposed method.

1. INTRODUCTION

Independent component analysis (ICA) [Hyvärinen et al., 2001] is a tool for statistical data analysis and signal processing that is able to recover independent input signals from mixed signals where the mixing process is unknown. This method has many potential applications in various fields such as speech signal processing, image processing biomedical signal processing and telecommunications. In the field of system control engineering, it has been applied to system identification [Sugimoto and Nitta, 2005], disturbance suppression [Sugimoto et al., 2005], fault detection [Sugimoto et al., 2005] and process monitoring [Kano et al., 1997].

Typical problems treated in ICA are called Blind Source Separation (BSS). While a standard BSS problem adopts a static mapping as the mixing process model, a blind deconvolution problem employs a dynamical equation instead. In system control engineering, we have to deal with dynamical systems, and hence have to solve the blind deconvolution problem. To solve this problem, an FIR filter approach [Thi and Jutten, 1995], [Gorokhov and Loubaton, 1999] and state-space approaches [Waheed and Salem, 2003], [Zhang and Cichocki, 2000], [Fukunaga and Fujimoto, 2006] were proposed. In these methods, the input signal is recovered by estimating the parameter of the inverse of the mixing system. However, if the system is nonminimum phase, the estimated parameter diverges due to the instability of the inverse system. For this problem, Zhang et al. proposed an ICA method for nonminimum phase systems using FIR filters [Zhang et al., 2004]. This method approximate the inverse system using FIR filters.

On the other hand, several researchers proposed stable inversion techniques for nonminimum phase systems [Devasia et al., 1996], [Hunt et al., 1996]. George et al. proposed a stable dynamic model inversion technique [Devasia et al., 1996], [Hunt et al., 1996]. This method construct an inverse filter using Kalman filters. We developed an ICA method for nonminimum phase systems using Kalman filters. However, since a blind deconvolution problem treats non-gaussian signals, we proposed an ICA method using H_∞ filters [Takaba and Katayama, 1994] which is not assumed to treat gaussian signals.

This paper is organized as follows. In Section 2, a design of H_∞ filters is presented. In Section 3, we propose an ICA method for nonminimum phase systems. First, an approximate inverse filter is constructed based on an H_∞ filter. Second, the parameters of the inverse filter is estimated by a steepest descent method. In Section 4, a numerical example is given. Finally, Section 5 concludes the paper.

2. DESIGN OF H_∞ FILTERS

Here, we refer to [Takaba and Katayama, 1994] for a design of H_∞ filters.

Consider the following linear time-varying system

$$x_{t+1} = F_t x_t + G_t w_t$$
$$y_t = H_t x_t + v_t$$

where $x_t \in \mathbb{R}^n$, $y_t \in \mathbb{R}^p$, $w_t \in \mathbb{R}^r$ and $v_t \in \mathbb{R}^p$ are the state vector, the measurement, the process disturbance and the measurement noise, respectively. Note that $w_t \in \mathbb{R}^r$ and $v_t \in \mathbb{R}^p$ are unknown and arbitrary $L_2[0, N]$ signals. We assume that an estimate of the initial state x_0 is given by $\hat{x}_0$. We also define

$$z_t = L_t x_t$$

where $z_t \in \mathbb{R}^q$, $L_t \in \mathbb{R}^{q\times n}$.

978-1-1234-7890-2/08/$20.00 © 2008 IFAC

10.3182/20080706-5-KR-1001.3941

The finite-horizon H_∞ filtering problem is to find the estimates $\hat{x}_t$ and $\hat{z}_t$ based on the measurement signal y_t so that the following inequality hold.

$$\sup_{x_0,v_t,w_t} \frac{\sum_{t=0}^{N} \|z_t - \hat{z}_t\|^2}{\|x_0 - \hat{x}_0\|^2_{\Sigma_0^{-1}} + \sum_{t=0}^{N} \|v_t\|^2_{R_t^{-1}} + \sum_{t=0}^{N} \|w_t\|^2_{Q_t^{-1}}} < \gamma^2 \tag{1}$$

where γ is a given constant and Σ_0, Q_t and R_t are positive definite weighting matrices.

We define the following cost function

$$\begin{aligned}\mathcal{J}(\hat{z}_t; x_0, v_t, w_t) := \sum_{t=0}^{N} \|z_t - \hat{z}_t\|^2 - \gamma^2 \Bigg(\|x_0 - \hat{x}_0\|^2_{\Sigma_0^{-1}} \\ + \sum_{t=0}^{N} \|v_t\|^2_{R_t^{-1}} + \sum_{t=0}^{N} \|w_t\|^2_{Q_t^{-1}} \Bigg)\end{aligned} \tag{2}$$

The inequality condition (1) is equivalent to having $\mathcal{J}(\hat{z}_t; x_0, v_t, w_t) < 0$. The problem is to estimate $\hat{z}_t$ satisfying the following inequality

$$\max_{x_0,v_t,w_t} \mathcal{J}(\hat{z}_t; x_0, v_t, w_t) < 0$$

Thus, we consider the following minimax problem between $\hat{z}_t$ and x_0, v_t, w_t.

$$\min_{\hat{z}_t} \max_{x_0,v_t,w_t} \mathcal{J}(\hat{z}_t; x_0, v_t, w_t) < 0$$

The optimal minimizer of the minimax problem is given by

$$\begin{aligned}\hat{z}_t &= L_t \hat{x}_{t|t} \\ \hat{x}_{t+1|t} &= F_t \hat{x}_{t|t} \\ \hat{x}_{t|t} &= F_t \hat{x}_{t|t-1} + K_t (y_t - H_t \hat{x}_{t|t-1})\end{aligned}$$

where

$$K_t := P_t H_t^{\mathrm{T}} (H_t P_t H_t^{\mathrm{T}} + R_t)^{-1}$$

$$P_{t+1} = F_t P_t \Psi_t^{-1} F_t^{\mathrm{T}} + G_t Q_t G_t^{\mathrm{T}} \tag{3}$$

$$\Psi_t = I + (H_t^{\mathrm{T}} R_t^{-1} H_t - \gamma^{-2} L_t^{\mathrm{T}} L_t) P_t \tag{4}$$

If γ is sufficiently large, the second term of equation (2) becomes dominant. Thus, as γ tends to infinity, the minimax problem reduces to the problem of minimizing

$$\begin{aligned}&\mathcal{J}(x_0, v_t, w_t) \\ &:= \|x_0 - \hat{x}_0\|^2_{\Sigma_0^{-1}} + \sum_{t=0}^{N} \|v_t\|^2_{R_t^{-1}} + \sum_{t=0}^{N} \|w_t\|^2_{Q_t^{-1}}\end{aligned}$$

with respect to w_t and x_t. It is well-known that this minimization problem is equivalent to the least mean square (LMS) estimation problem in the case where x_0 is generated by the gaussian distribution $\mathcal{N}(\hat{x}_0, \Sigma_0)$ and where w_t and v_t are the zero mean gaussian white noises with unit covariances. The optimal solution of the LMS estimation problem is given by the Kalman filter.

3. INDEPENDENT COMPONENT ANALYSIS FOR NONMINIMUM PHASE SYSTEMS

3.1 Problem setting

Consider the following linear time-invariant system

$$\begin{aligned}x_{t+1} &= A x_t + B u_t \\ y_t &= C x_t + D u_t\end{aligned} \tag{5}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$ and $y_t \in \mathbb{R}^m$ are the state vector of the system, the input signal and the measured signal, respectively. Here we make the following assumption for the input.

Assumption 1. The elements of the unknown input signal u_t are stationary zero-mean i.i.d. process and mutually statistically independent.

Here, y_t is known and A, B, C, D, x_t and u_t are known. The objective is to estimate the input u_t only from the output y_t. This can be achieved by finding the estimates $\hat{A}_t$, $\hat{B}_t$, $\hat{C}_t$, $\hat{D}_t$ and $\hat{x}_t$ of A, B, C, D and x_t.

3.2 Design of an approximate inverse filter

In the basic ICA approach, the input signal is recovered by estimating the parameter of the inverse of the mixing system. However, if the system is nonminimum phase, the estimated parameter diverges due to the instability of the inverse. This subsection constructs an approximate inverse filter using H_∞ filters

In the stable dynamic inversion technique, the inverse filter is constructed based on the Kalman filter. We design the inverse filter using the H_∞ filter. Since A, B, C and D are unknown, we consider the following time-varying systems

$$\begin{aligned}x_{t+1} &= \hat{A}_t x_t + \hat{B}_t u_t \\ y_t &= \hat{C}_t x_t + \hat{D}_t u_t\end{aligned} \tag{6}$$

For this system, the H_∞ filter is constructed by

$$\begin{aligned}\hat{x}_{t+1|t} &= \hat{A}_t \hat{x}_t \\ \hat{x}_{t|t} &= \hat{x}_{t|t} - K_t (\hat{D}_t^{-1} \hat{C}_t \hat{x}_{t|t} - \hat{D}_t^{-1} y_t)\end{aligned} \tag{7}$$

where

$$K_t := P_t (\hat{D}_t^{-1} \hat{C}_t)^{\mathrm{T}} (\hat{D}_t^{-1} \hat{C}_t P_t (\hat{D}_t^{-1} \hat{C}_t)^{\mathrm{T}} + R_t)^{-1} \tag{8}$$

$$P_{t+1} = \hat{A}_t P_t \Psi_t^{-1} \hat{A}_t^{\mathrm{T}} + \hat{B}_t Q_t \hat{B}_t^{\mathrm{T}} \tag{9}$$

$$\Psi_t = I + ((\hat{D}_t^{-1} \hat{C}_t)^{\mathrm{T}} R_t^{-1} \hat{D}_t^{-1} \hat{C}_t - \gamma^{-2} L_t^{\mathrm{T}} L_t) P_t \tag{10}$$

The input signal can be estimated by using the inverse of the output equation (6). Therefore, the inverse filter is given by

$$\begin{aligned}\hat{x}_{t+1|t} &= \hat{A}_{s,t}\, \hat{x}_{t|t-1} + \hat{A}_t K_t \hat{\hat{D}}_t y_t \\ \hat{u}_t &= -\hat{\hat{C}}_t \hat{x}_{t|t-1} + \hat{\hat{D}}_t y_t\end{aligned} \tag{11}$$

where

$$\begin{aligned}\hat{A}_{s,t} &= \hat{A}_t - \hat{A}_t K_t \hat{\hat{C}}_t \\ \hat{\hat{C}}_t &= \hat{D}_t^{-1} \hat{C}_t \\ \hat{\hat{D}}_t &= \hat{D}_t^{-1}\end{aligned}$$

If the matrix D is singular, $\hat{D}_t^{-1}\hat{C}_t$ and $\hat{D}_t^{-1}$ cannot be calculate. We estimate $\hat{\hat{C}}_t$ and $\hat{\hat{D}}_t$ in order to avoid a singular point.

3.3 Estimation of the parameter of the inverse filter

In this subsection, we derive the algorithm for estimating the parameters $\hat{A}_t$, $\hat{B}_t$, $\hat{\hat{C}}_t$, $\hat{\hat{D}}_t$ of the inverse filter.

A densed notation of the outputs $\mathcal{Y}$ and that of the estimates of the inputs $\hat{\mathcal{U}}$ are defined by

$$\mathcal{Y} := (y_0^{\mathrm{T}}, y_1^{\mathrm{T}}, ..., y_N^{\mathrm{T}})^{\mathrm{T}}$$
$$\hat{\mathcal{U}} := (\hat{u}_0^{\mathrm{T}}, \hat{u}_1^{\mathrm{T}}, ..., \hat{u}_N^{\mathrm{T}})^{\mathrm{T}}$$

If the estimates of the inputs $\hat{\mathcal{U}}$ are spatially mutually independent and temporarily i.i.d. signals, then the probability density function $p(\hat{\mathcal{U}})$ of $\hat{\mathcal{U}}$ is factorized as follows:

$$p(\hat{\mathcal{U}}) \equiv \prod_{i=1}^{m}\prod_{t=0}^{N} p(\hat{u}_{i,t})$$

Statistical independence can be measured using Kullback-Leibler divergence between $p(\hat{\mathcal{U}})$ and $\prod_{i=1}^{m}\prod_{t=0}^{N} p(\hat{u}_{i,t})$

$$I(\hat{\mathcal{U}}) := \int p(\hat{\mathcal{U}}) \log \frac{p(\hat{\mathcal{U}})}{\prod_{i=1}^{m}\prod_{t=0}^{N} p(\hat{u}_{i,t})} \mathrm{d}\hat{\mathcal{U}} \tag{12}$$

The objective is to minimize $I(\hat{\mathcal{U}})$ with respect to the parameters $\hat{A}_t$, $\hat{B}_t$, $\hat{\hat{C}}_t$ and $\hat{\hat{D}}_t$ by the steepest descent method.

We reformulate the cost function in order to implement the on-line learning for the parameters of the inverse filter.

According to the property of the probability density function, we derive the following relation between $p(\hat{\mathcal{U}})$ and $p(\mathcal{Y})$:

$$p(\hat{\mathcal{U}}) = \frac{p(\mathcal{Y})}{|J|} \tag{13}$$

where J is the determinant of the Jacobian matrix which is calculated as follows

$$J = \det \begin{pmatrix} \hat{\hat{D}}_0 & 0 & \dots & 0 \\ * & \hat{\hat{D}}_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & \hat{\hat{D}}_N \end{pmatrix}$$

where $*$ denotes a nonzero value. Since the above Jacobian matrix is lower block triangular, $p(\hat{\mathcal{U}})$ is described by

$$p(\hat{\mathcal{U}}) = \frac{p(\mathcal{Y})}{\prod_{t=0}^{N} |\det \hat{\hat{D}}_t|} \tag{14}$$

Substituting equation (14) for equation (12) gives

$$I(\hat{\mathcal{U}}) = -\sum_{t=0}^{N} E[\log|\det \hat{\hat{D}}_t|] + E[\log p(\mathcal{Y})] - \sum_{i=1}^{m}\sum_{t=0}^{N} E[\log p(\hat{u}_{i,t})]$$

Here $E[\log p(\mathcal{Y})]$ is a constant. Therefore, the cost function can be simplified as

$$\tilde{I}(\hat{u}_t) := -\log|\det \hat{\hat{D}}_t| - \sum_{i=1}^{m} \log p(\hat{u}_{i,t}) \tag{15}$$

The minimization of $\tilde{I}(\hat{u}_t)$ requires the computation of its gradient with respect to the parameters $\hat{A}_t$, $\hat{B}_t$, $\hat{\hat{C}}_t$ and $\hat{\hat{D}}_t$. The gradients are computed separately in the parameters of the output equation and those of the dynamics.

We can calculate the derivative of $\tilde{I}(\hat{u}_t)$ with respect to the matrices $\hat{\hat{C}}_t$ and $\hat{\hat{D}}_t$ as

$$\frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{\hat{C}}_t} = -\varphi(\hat{u}_t)\hat{x}_{t|t-1}^{\mathrm{T}} \tag{16}$$

$$\frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{\hat{D}}_t} = -\hat{\hat{D}}_t^{-\mathrm{T}} + \varphi(\hat{u}_t)y_t^{\mathrm{T}} \tag{17}$$

where $\varphi(\hat{u}_t)$ is defined by

$$\varphi(\hat{u}_t) := -\left(\frac{\mathrm{d}\log p(\hat{u}_{1,t})}{\mathrm{d}\hat{u}_{1,t}}, ..., \frac{\mathrm{d}\log p(\hat{u}_{m,t})}{\mathrm{d}\hat{u}_{m,t}}\right)^{\mathrm{T}} \tag{18}$$

The nonlinear function $\varphi(\hat{u})$ depends on the probability density function of the input signal, which is unknown in a blind deconvolution setting. It is not necessary to estimate the probability density function precisely. One important issue in determining the nonlinear function is that the stability conditions of the learning algorithm must be satisfied [Amari et al., 1997]. The gradient of the cost function in (17) requires calculating the inverse of the matrix $\hat{\hat{D}}_t$. Fortunately, a modification of standard gradient-based procedures has been developed that overcomes this difficulty in the BSS task. This modification, termed the natural gradient by [Amari et al., 1996], modifies the standard gradient update by a linear transformation whose elements are determined by the Riemannian metric tensor for the assumed parameter space. Then the modified search direction is given by

$$\frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{\hat{D}}_t}\hat{\hat{D}}_t^{\mathrm{T}}\hat{\hat{D}}_t = -(I - \varphi(\hat{u}_t)y_t^{\mathrm{T}}\hat{\hat{D}}_t^{\mathrm{T}})\hat{\hat{D}}_t \tag{19}$$

The gradients of $\tilde{I}(\hat{u}_t)$ with respect to the matrices $\hat{A}_t$ and $\hat{B}_t$ can be calculated as

$$\frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{A}_{t-1}} = \sum_{p=1}^{n} \frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{x}_{p,t|t-1}} \frac{\partial \hat{x}_{p,t|t-1}}{\partial \hat{A}_{t-1}} \tag{20}$$

$$\frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{B}_{t-2}} = \sum_{p=1}^{n} \frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{x}_{p,t|t-1}} \frac{\partial \hat{x}_{p,t|t-1}}{\partial \hat{B}_{t-2}} \tag{21}$$

(See the Appendix for the details of this calculation.)

3.4 Algorithm

Summarizing the results of the previous section, the block diagram of the proposed method is shown in Fig 1 and the proposed algorithm consists of the following steps:
Algorithm 1.

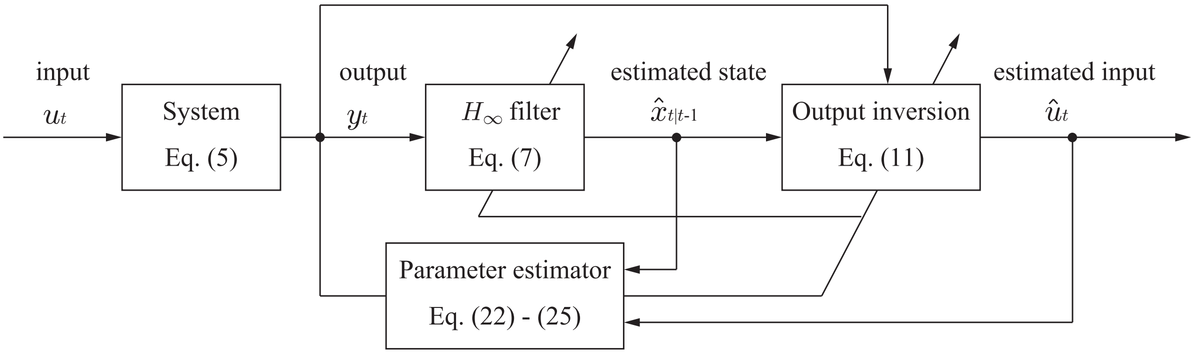

Fig. 1. Block diagram of the proposed method

step1 Set the initial value of the parameters $\hat{A}_0$, $\hat{B}_0$, $\hat{\bar{C}}_0$, $\hat{\bar{D}}_0$ and P_0

step2 Compute the gain (8)

step3 Update the filter equation (7)

step4 Update the matrix (9)

step5 Update the parameters $\hat{A}_t$, $\hat{B}_t$, $\hat{\bar{C}}_t$, $\hat{\bar{D}}_t$ using the gradients are described by equations (20), (21), (16) and (19)

$$\hat{A}_{t+1} = \hat{A}_t + \mu_1 \sum_{p=1}^{n} \frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{x}_{p,t|t-1}} \frac{\partial \hat{x}_{p,t|t-1}}{\partial \hat{A}_{t-1}} \tag{22}$$

$$\hat{B}_{t+1} = \hat{B}_t + \mu_2 \sum_{p=1}^{n} \frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{x}_{p,t|t-1}} \frac{\partial \hat{x}_{p,t|t-1}}{\partial \hat{B}_{t-2}} \tag{23}$$

$$\hat{\bar{C}}_{t+1} = \hat{\bar{C}}_t + \mu_3 \varphi(\hat{u}_t) \hat{x}_{t|t-1}^{\mathrm{T}} \tag{24}$$

$$\hat{\bar{D}}_{t+1} = \hat{\bar{D}}_t + \mu_4 (I - \varphi(\hat{u}_t) y_t^{\mathrm{T}} \hat{\bar{D}}_t^{\mathrm{T}}) \hat{\bar{D}}_t \tag{25}$$

where $\mu_1, \mu_2, \mu_3, \mu_4 > 0$ are the step size parameters.

step6 Compute the estimate of the input (11)

4. NUMERICAL EXAMPLE

The parameters of the system (5) are

$$A = \begin{pmatrix} 0 & 0 & 1.00 & 0 \\ 0 & 0 & 0 & 1.00 \\ 0.5 & -0.41 & -0.70 & 0.72 \\ 0.72 & -0.72 & -0.43 & 0.17 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$C = \begin{pmatrix} -0.05 & -0.46 & -0.85 & 1.13 \\ -0.19 & -0.24 & -0.99 & -0.47 \end{pmatrix}$$

$$D = \begin{pmatrix} 0.84 & -0.66 \\ 0.68 & 0.12 \end{pmatrix}$$

The zeros are at -0.8651, 1.3877, 0.8458, $-0.0830 \pm 1.1228i$, $0.2506 \pm 0.7894i$, -0.2957, 0.8458, -0.6776, $0.7059 \pm 0.9828i$, 0.4479, 5.2382, 0.5003, -2.0528 and hence the system is nonminimum phase. The input signal u_t is chosen to be i.i.d signal uniformly distributed in the range $(-1, 1)$.

The nonlinear function is chosen as $\varphi(\hat{u}_t) = \hat{u}_t^3$. The initial parameters are determined as follows:

$$\hat{A}_0 = \begin{pmatrix} 0.0698 & 0.1663 & 1.0770 & 0.2460 \\ 0.0832 & 0.2566 & 0.2777 & 1.1302 \\ 0.5751 & -0.1129 & -0.4923 & 0.7588 \\ 0.7254 & -0.5943 & -0.2498 & 0.2921 \end{pmatrix}$$

$$\hat{B}_0 = \begin{pmatrix} 0.1616 & 0.2454 \\ 0.2444 & 0.0616 \\ 1.2197 & 0.2448 \\ 0.2655 & 1.1193 \end{pmatrix}$$

$$\hat{\bar{C}}_0 = \begin{pmatrix} 0.2126 & -0.1681 & -1.1946 & 0.0632 \\ 0.0193 & 0.3598 & -0.2572 & -1.6767 \end{pmatrix}$$

$$\hat{\bar{D}}_0 = \begin{pmatrix} 0.6180 & 1.2460 \\ -0.8785 & 1.9261 \end{pmatrix}$$

Fig. 2 shows the source signals and the recovered signals. In the figure, the solid line denote the recovered signal and the dashed line denote the input signal. Fig. 3 shows the cost function $\tilde{I}(\hat{u}_t)$. The computation of the cost function needs to use the probability density function of $\hat{u}_t$ which is unknown. We solve equation (18) with nonlinear function $\varphi(\hat{u}_t) = \hat{u}_t^3$. The solution of equation (18) is given by

$$p(\hat{u}_t) = \frac{\sqrt{2}}{\Gamma\left(\frac{1}{4}\right)} \exp\left(-\frac{\hat{u}_t^4}{4}\right)$$

where $\Gamma(\cdot)$ is the gamma function. We observed that the value of the cost function decreases and converges. This result demonstrates the effectiveness of the proposed method.

To illustrate the influence of γ, we used $\gamma = 1.15$ and ∞ Fig. 4 shows the cost function averaged over 20 steps. In the figure, the solid line denotes $\gamma = 1.15$ and the dashed line denotes $\gamma = \infty$. The error bar denotes the standard derivation from the average of 10 trials. The value of the cost function at $\gamma = 1.15$ is smaller than $\gamma = \infty$.

We compare the performance of the proposed method and the Zhang's method [Zhang et al., 2004]. The filter length of the Zhang's method is 10. Fig. 5 shows the cost function averaged over 20 steps. In the figure, the solid line denotes the proposed method and the dashed-dotted line denotes the Zhang's method. The error bar denotes the standard derivation from the average of 10 trials. The proposed method converges faster than the Zhang's method. Furthermore it is seen that the proposed method can obtain a smaller cost function than the Zhang's method.

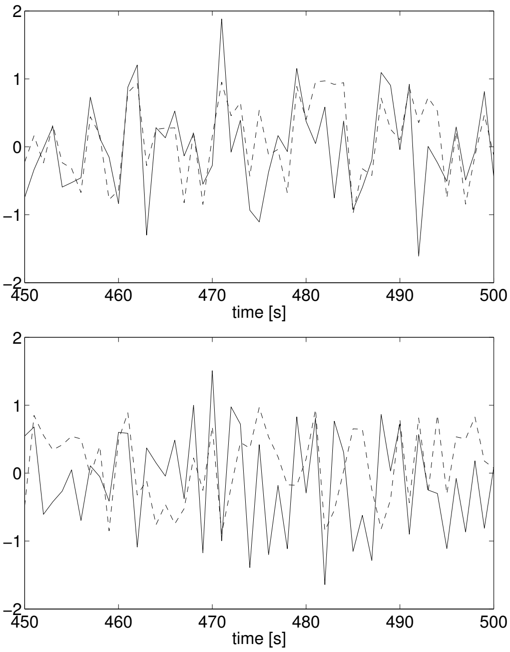

Fig. 2. Source signals and recovered signals

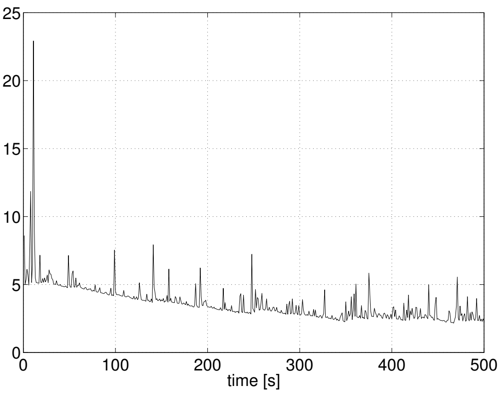

Fig. 3. Cost function $\tilde{I}(\hat{u}_t)$

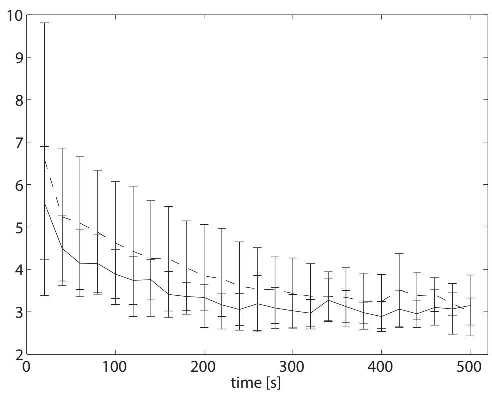

Fig. 4. Cost function in case of $\gamma = 1.15$ and ∞

5. CONCLUSION

In this paper, an independent component analysis method for nonminimum phase systems using H_∞ filters is proposed. An inverse filter is constructed based on an H_∞ filter in order to estimate the state of the given plant.

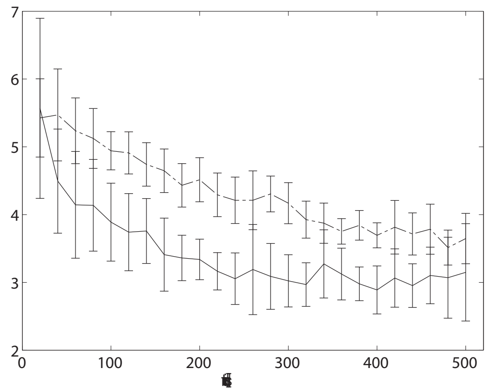

Fig. 5. Cost function of the proposed method and the Zhang's method

The learning algorithm to estimate the parameter of the system is derived by minimizing the Kullback-Leibler divergence. A numerical simulation shows the effectiveness of the proposed method.

REFERENCES

S. Amari, T.P. Chen, and A. Cichocki. Stability analysis of learning algorithms for blind source separation. *Neural Networks*, 10(8):1345–1351, 1997.

S. Amari, A. Cichocki, and H. H. Yang. A new learning algorithm for blind signal separation. In *Advances in Neural Information Processing Systems 8*, pages 757–763, 1996.

S. Devasia, D. Chen, and B. Paden. Nonlinear inversion-based output tracking. *IEEE Transactions on Automatic Control*, 41(7):930–942, 1996.

S. Fukunaga and K. Fujimoto. Nonlinear blind deconvolution based on a state-space model. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 6259–6300, 2006.

A. Gorokhov and P. Loubaton. Blind identification of MIMO-FIR systems: a generalized linear prediction approach. *Signal Processing*, 73(1-2):105–124, 1999. ISSN 0165-1684.

L.R. Hunt, G. Meyer, and R. Su. Noncausal inverses for linear systems. *IEEE Transactions on Automatic Control*, 41(4):608–611, 1996.

A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, 2001.

M. Kano, S. Tanaka, H. Maruta, S. Hasebe, I. Hashimoto, and H. Ohno. Statistical process monitoring with external analysis and independent component analysis. *Transactions of the Society of Instrument and Control Engineers*, 38(11), 1997. (in Japanese).

K. Sugimoto and M. Nitta. Polynomial matrix approach to independent component analysis: (part i) basics. In *IFAC World Congress*, 2005.

K. Sugimoto, A. Suzuki, M. Nitta, and N. Adachi. Polynomial matrix approach to independent component analysis: (part ii) application. In *IFAC World Congress*, 2005.

K. Takaba and T. Katayama. Finite-horizon minimax estimation for linear discrete-time systems. *Transactions of Institute of Systems, Control and Information Engineers*, 7(8):322–331, 1994. (in Japanese).

H.L. Nguyen Thi and C. Jutten. Blind source separation for convolutive mixtures. *Signal processing*, 45:209–229, 1995.

K. Waheed and F.M. Salem. Blind source recovery: A framework in the state space. *Journal of Machine Learning Research*, 4:1411–1446, 2003.

L. Zhang and A. Cichocki. Blind deconvolution of dynamical systems: A state space approach. *Journal of Signal Processing*, 4(2):111–130, 2000.

L. Zhang, A. Cichocki, and S. Amari. Multichannel blind deconvolution of nonminimum-phase systems using filter decomposition. *IEEE Transactions on Signal Processing*, 52(5):1430–1442, 2004.

Appendix A. CALCULATION OF THE GRADIENTS WITH RESPECT TO THE PARAMETERS $\hat{A}$ AND $\hat{B}$

The gradients with respect to the parameters $\hat{A}_t$ and $\hat{B}_t$ are calculated by

$$\frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{A}_{ij,t-1}} = -\varphi(\hat{u}_t)^{\mathrm{T}} \sum_{p=1}^{n} \hat{\hat{C}}_{p,t} \frac{\partial \hat{x}_{p,t|t-1}}{\partial \hat{A}_{ij,t-1}}$$

$$\frac{\partial \tilde{I}(\hat{u}_t)}{\partial \hat{B}_{ik,t-2}} = -\varphi(\hat{u}_t)^{\mathrm{T}} \sum_{p=1}^{n} \hat{\hat{C}}_{p,t} \frac{\partial \hat{x}_{p,t|t-1}}{\partial \hat{B}_{ik,t-2}}$$

where $\hat{\hat{C}}_p$ is the l-th column vector of matrix $\hat{\hat{C}}$. Here, $\partial \hat{x}_{p,t|t-1}/\partial \hat{A}_{ij,t-1}$ and $\partial \hat{x}_{p,t|t-1}/\partial \hat{B}_{ik,t-2}$ are described by

$$\begin{aligned}\frac{\partial \hat{x}_{t+1|t}}{\partial \hat{A}_{ij,t}} &= \hat{A}_{s,t} \frac{\partial \hat{x}_{t|t-1}}{\partial \hat{A}_{ij,t}} + J^{ij} \hat{x}_{t|t-1} \\ &\quad - J^{ij} K_t \left(\hat{\hat{C}}_t \hat{x}_{t|t-1} - \hat{\hat{D}}_t y_t \right) \\ \frac{\partial \hat{x}_{t+1}}{\partial \hat{B}_{ik,t-1}} &= \hat{A}_{s,t} \frac{\partial \hat{x}_{r,t}}{\partial \hat{B}_{ik,t-1}} \\ &\quad - \hat{A}_t \frac{\partial K_t}{\partial \hat{B}_{ik,t-1}} \left(\hat{\hat{C}}_t \hat{x}_t - \hat{\hat{D}}_t y_t \right)\end{aligned}$$

where J^{ij} is the single-entry matrix with 1 at (i, j) and zero elsewhere. $\partial K_t/\partial \hat{B}_{ik,t}$ is given by

$$\begin{aligned}\frac{\partial K_t}{\partial \hat{B}_{ik,t-1}} &= \frac{\partial P_t}{\partial \hat{B}_{ik,t-1}} \hat{\hat{C}}_t^{\mathrm{T}} (\hat{\hat{C}}_t P_t \hat{\hat{C}}_t^{\mathrm{T}} + R_t)^{-1} \\ &\quad - P_t \hat{\hat{C}}_t^{\mathrm{T}} (\hat{\hat{C}}_t P_t \hat{\hat{C}}_t^{\mathrm{T}} + R_t)^{-1} \\ &\quad \times \hat{\hat{C}}_t \frac{\partial P_t}{\partial \hat{B}_{ik,t-1}} \hat{\hat{C}}_t^{\mathrm{T}} (\hat{\hat{C}}_t P_t \hat{\hat{C}}_t^{\mathrm{T}} + R_t)^{-1} \\ \frac{\partial P_{t+1}}{\partial \hat{B}_{ik,t}} &= J^{ik} Q_t \hat{B}_t^{\mathrm{T}} + \hat{B}_t Q_t {J^{ik}}^{\mathrm{T}}\end{aligned}$$