Analysis of signal separation and signal distortion in feedforward and feedback blind source separation based on source spectra

メタデータ	言語: English 出版者: 公開日: 2017-10-03 キーワード (Ja): キーワード (En): 作成者: Horita, Akihide, Nakayama, Kenji, Hirano, Akihiro, Dejima, Yasuhiro メールアドレス: 所属:
URL	http://hdl.handle.net/2297/6815

Analysis of Signal Separation and Signal Distortion in Feedforward and Feedback Blind Source Separation Based on Source Spectra

Akihide Horita Kenji Nakayama Akihiro Hirano Yasuhiro Dejima
Graduate School of Natural Science and Technology, Kanazawa Univ.
Kakuma-machi, Kanazawa, 920-1192, Japan
E-mail: horita@leo.ec.t.kanazawa-u.ac.jp, nakayama@t.kanazawa-u.ac.jp

***Abstract*— Source separation and signal distortion in three kinds of BSSs with convolutive mixture are analyzed. They include a feedforward BSS, trained in the time domain and in the frequency domain, and a feedback BSS, trained in the time domain. First, an evaluation measure of signal distortion is discussed. Second, conditions for source separation and distortion free are derived. Based on these conditions, source separation and signal distortion are analyzed. The feedforward BSS has some degree of freedom, and the output spectrum can be changed. The feedforward BSS, trained in the frequency domain, has weighting effect, which can suppress signal distortion. This weighting is, however, effective only when the source spectra are similar to each other. Since, the feedforward BSS, trained in the time domain, does not have any constraints on signal distortion free, its output signals can be easily distorted. A new learning algorithm with a distortion free constraint is proposed. On the other hand, the feedback BSS can satisfy both source separation and distortion free conditions simultaneously. Simulation results support the theoretical analysis.**

I. Introduction

Since, in many applications, mixing processes are convolutive mixtures, several methods in the time domain and the frequency domain have been proposed. Two kinds of network structures have been proposed, including feedforward (FF) and feedback (FB) structures. Separation performance is highly dependent on the signal sources and the transfer functions in the mixture [5]-[8],[11],[12],[14],[15].

The BSS learning algorithms make the output signals to be statistically independent. This direction cannot always guarantee distortion free separation. Some signal distortion may be caused. A regularization method has been proposed, in which the distance between the observed signals and the separated signals is added to the cost function. However, since the observations include many kinds of signal sources, it is difficult to suppress signal distortion. Furthermore, even though the signal distortion in the BSSs is an important problem, it has not been well discussed [16].

In this paper, first, an evaluation measure of signal distortion is discussed. Second, conditions for source separation and signal distortion free are derived. Based on these conditions, convergence properties are analyzed. A new learning algorithm for the FF-BSS, trained in the time domain, is also proposed. Finally, simulation results will be demonstrated in order to confirm results of the theoretical analysis.

II. FF-BSS for Convolutive Mixture

A. Network Structure and Equations

For simplicity, 2 signal sources and 2 sensors are used. A block diagram is shown in Fig.1. The observations and the

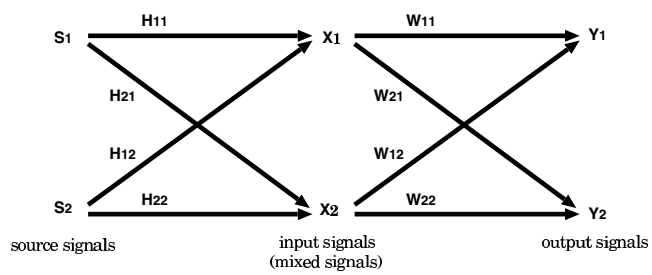

Fig. 1. FF-BSS with 2 signal sources and 2 sensors.

output signals are given by

$$x_j(n) = \sum_{i=1}^{2}\sum_{l=0}^{K_h-1} h_{ji}(l)s_i(n-l), j=1,2 \quad (1)$$

$$y_k(n) = \sum_{j=1}^{2}\sum_{l=0}^{K_w-1} w_{kj}(l)x_j(n-l), k=1,2 \quad (2)$$

B. Learning Algorithm in Time Domain

The learning algorithm is derived following the natural gradient algorithm using the mutual information as a cost function [4].

$$w_{kj}(n+1,l) = w_{kj}(n,l) + \eta\{w_{kj}(n,l) - \sum_{\substack{p=1\\ \neq j}}^{2}\sum_{q=0}^{K_w-1} \varphi(y_k(n))y_p(n-l+q)w_{pj}(n,q)\} \quad (3)$$

$$\varphi(y_k(n)) = \frac{1-e^{-y_k(n)}}{1+e^{-y_k(n)}} \quad (4)$$

η is a learning rate.

C. Learning Algorithm in Frequency Domain

Filter coefficients in the separation block are trained according [4],[9],[10],[13],

$$\boldsymbol{W}(r+1,m) = \boldsymbol{W}(r,m) + \eta[diag(\langle\Phi(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle - \langle\Phi(\boldsymbol{Y}(r,m))\boldsymbol{Y}^H(r,m)\rangle]\boldsymbol{W}(r,m) \quad (5)$$

$$\Phi(\boldsymbol{Y}(r,m)) = \frac{1}{1+e^{-\boldsymbol{Y}^R(r,m)}} + \frac{j}{1+e^{-\boldsymbol{Y}^I(r,m)}} \quad (6)$$

r is the block number used in FFT, and m indicates the frequency point in each block. $<>$ is an averaging operation. $\boldsymbol{W}(r,m)$ is a weight matrix of the r-th block FFT and the m-th frequency point. Its (k,j) element is $W_{kj}(r,m)$, which is the connection from the j-th observation to the k-th output. $\boldsymbol{Y}(r,m)$ is the output vector of the r-th block FFT and the m-th frequency point. Its k-th element is $Y_k(r,m)$, which is the k-th output. $\boldsymbol{Y}^R(r,m)$ and $\boldsymbol{Y}^I(r,m)$ indicate the real part and the imaginary part.

III. FB-BSS for Convolutive Mixture

A. Network Structure and Equations

Figure 2 shows an FB-BSS proposed by Jutten et all [1]. The mixing stage has a convolutive structure. C_{ij} consists of an FIR filter.

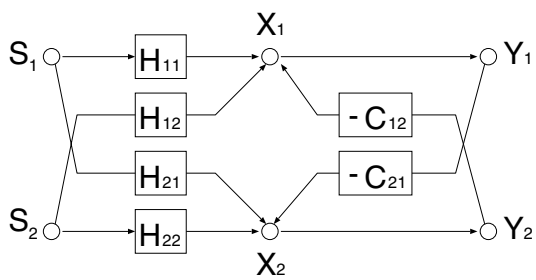

Fig. 2. FB-BSS with 2 signal sources and 2 sensors.

The observations and the output signals are expressed as follows:

$$x_j(n) = \sum_{i=1}^{N}\sum_{m=0}^{M_{ji}-1} h_{ji}(m)s_i(n-m) \quad (7)$$

$$y_k(n) = x_k(n) - \sum_{\substack{k=1\\ \neq j}}^{N}\sum_{l=0}^{L_{jk}-1} c_{jk}(l)y_k(n-l) \quad (8)$$

B. Learning Algorithm

In the FB-BS, the learning algorithm in the time domain is used [3]. The following learning algorithm has been derived assuming some conditions [12],[14]. The signal sources $S_1(z)$ and $S_2(z)$ locate close to the sensors of $X_1(z)$ and $X_2(z)$, respectively. Therefore, time delay of $H_{ji}(z), i \neq j$ are slightly longer than those of $H_{ii}(z)$. Furthermore, amplitude responses of $H_{ji}(z), i \neq j$ are smaller than those of $H_{ii}(z)$. These conditions are practically acceptable.

$$\begin{aligned} c_{jk}(n+1,l) &= c_{jk}(n,l) \\ &+ \eta f(y_j(n))g(y_k(n-l)) \end{aligned} \quad (9)$$

$f(y_j(n))$ and $g(y_k(n-l))$ are odd functions.

IV. Criterion of Signal Distortion

How to evaluate signal distortion in the BSSs is one of the problems. The learning algorithms used in the BSS make the output signals to be statistically independent. Estimation of the mixing process is not taken into account. Especially, in the convolutive mixtures, the output signals are not guaranteed to approach to the sources. Therefore, the signal sources observed at the sensors are taken into account as a criterion for the signal distortion [3],[16].

In this paper, the signal distortion is evaluated as a distance from the observed signal sources. However, in this case, several criteria can be considered. The signal sources included in the observations $x_j(n)$ are given by $H_{ii}(z)S_i(z)$ and $H_{ji}(z)S_i(z), i \neq j$. How to combine these components will provide several criteria. The following measures are considered.

$$\begin{aligned} \sigma_{d1a} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})S_i(e^{j\omega}) \\ &- Y_k(e^{j\omega})|^2 d\omega \end{aligned} \quad (10)$$

$$\begin{aligned} \sigma_{d1b} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} (|H_{ji}(e^{j\omega})S_i(e^{j\omega})| \\ &- |Y_k(e^{j\omega})|)^2 d\omega \end{aligned} \quad (11)$$

$$\sigma_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (12)$$

$$SD_{1x} = 10\log_{10}\frac{\sigma_{d1x}}{\sigma_1}, x = a, b \quad (13)$$

$$\sigma_{d2a} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega}) - A_{ki}(e^{j\omega})|^2 d\omega \quad (14)$$

$$\sigma_{d2b} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (|H_{ji}(e^{j\omega})| - |A_{ki}(e^{j\omega})|)^2 d\omega \quad (15)$$

$$\sigma_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})|^2 d\omega \quad (16)$$

$$SD_{2x} = 10\log_{10}\frac{\sigma_{d2x}}{\sigma_2}, x = a, b \quad (17)$$

V. Source Separation and Signal Distortion in FF-BSS

A. Learning in Frequency Domain

For simplicity, the FF-BSS with 2-sources and 2-sensors, shown in Fig.1, is used. Furthermore, $S_i(z)$ is assumed to be separated at the output $Y_i(z)$. This does not lose generality. Taking the signal distortion criterion into account, the condition on distortion-free source separation can be expressed as follows:

$$W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) = H_{11}(z) \quad (18)$$

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (19)$$

$$W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) = 0 \quad (20)$$

$$W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z) = H_{22}(z) \quad (21)$$

The above equations include two kinds of conditions. One of them is the complete source separation, that is the non-diagonal elements are zero as shown in Eqs.(19) and (20). The other is the signal distortion free, that is the diagonal elements are $H_{ii}(z)$ as shown in Eqs.(18) and (21). These conditions are further modified.

From the relations of Eqs.(19) and (20), $H_{ji}(z)$ are ex-

pressed as follows:

$$H_{12}(z) = -\frac{W_{12}(z)}{W_{11}(z)}H_{22}(z) \quad (22)$$

$$H_{21}(z) = -\frac{W_{21}(z)}{W_{22}(z)}H_{11}(z) \quad (23)$$

By substituting the above equations into the relations of Eqs.(18) and (21), $H_{ji}(z)$ can be cancelled, and the following equations for only $W_{kj}(z)$ are obtained.

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{22}(z) \quad (24)$$

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{11}(z) \quad (25)$$

From these equations, $W_{11}(z) = W_{22}(z)$ is derived. Therefore, the above equations result in

$$W_{jj}^2(z) - W_{jj}(z) - W_{jk}(z)W_{kj}(z) = 0 \quad (26)$$
$$j = 1,2, k = 1,2, j \neq k$$

This 2nd-order equation expresses the condition on complete source separation without signal distortion. This equation is solved for $W_{11}(z)$ and $W_{22}(z)$ as follows:

$$W_{jj}(z) = \frac{1 \pm \sqrt{1 + 4W_{12}(z)W_{21}(z)}}{2}, j = 1,2 \quad (27)$$

This constraint can be included in the learning processes of the FF-BSS in the time domain and in the frequency domain.

B. Learning Algorithm with Constraint in Time Domain

The conventional learning algorithm given by Eqs.(3),(4) does not satisfy the condition given by Eq.(27). Usually, only Eqs.(19) and (20) are approximately satisfyed. Equations (18) and (21) are not guaranteed. Therefore, in general, signal distortion cannot be supressed.

In this section, a new learning algorithm for the FF-BSS, trained in the time domain, is proposed. The constraint given by Eq.(27) is taken into account in the learning process. Equation (27) is rewritten as follows:

$$(2W_{jj}(z) - 1)^2 = 1 + 4W_{12}(z)W_{21}(z) \quad (28)$$

This constraint is used in the learning process as follows: Given $W_{12}(z)$ and $W_{21}(z)$, the coefficients of $W_{jj}(z)$ are obtained so as to approximate the relation of Eq.(28).

The condition for the distortion free source separation is derived based on the complete separation and signal distortion free. However, the learning of the separation block begins from some initial guess. Therefore, in the early stage of the learning process, the signal sources are not well separated. Taking these situations into account, the constraint of Eq.(28) is gradually imposed as the learning process makes progress. The following learning algorithm is proposed.

$$w_{kj}(n+1,l) = w_{kj}(n) + \eta\{w_{kj}(n) - \sum_{o=0}^{K_w-1}\sum_{p=1}^{2}\phi(y_k(n))y_p(n-o+p)w_{kp}(n,o)\} \quad (29)$$

$$w_{jj}(N+1,l) = (1-\alpha)w_{jj}(n+1,l) + \alpha\tilde{w}_{jj}(n+1) \quad (30)$$

$\tilde{w}_{jj}(n+1)$ is determined so as to approximate the relation of Eq.(28). α is usually set to a small positive number.

C. Signal Distortion in FF-BSS Trained in Frequency Domain

The constraint given by Eq.(27) can be applied to the learning process in the frequency domain. Given $W_{12}(e^{j\omega})$ and $W_{21}(e^{j\omega})$, $W_{jj}(e^{j\omega})$ are calculated by Eq.(27) exactly. In this case, the constraint should be gradually imposed as described in Sec.V-B.

On the other hand, in the frequency domain, there is some weighting effect. From Eq.(5), the correction of the weights is highly dependent on $\boldsymbol{Y}$, that is the frequency response of the outputs. If the initial guess of $\boldsymbol{W}(r,m)$ is set to the identity matrix, that is $\boldsymbol{W}(0,m) = \boldsymbol{I}$, then $\boldsymbol{Y}(0,m) = \boldsymbol{X}(0,m)$, where $\boldsymbol{X}(r,m) = \boldsymbol{HS}(r,m)$. Therefore, the correction of $\boldsymbol{W}(r,m)$ is proportional to $\boldsymbol{HS}(r,m)$. If the signal sources are all speech, their spectra are similar to each other. In this case, the spectra of X_i, which are the composite signals of the signal sources, are also similar to those of speeches. This means the correction of $\boldsymbol{W}(r,m)$ is weighted by the spectra of the observed signal sources. Furthermore, as the learning makes progress, since Y_k gradually approaches to the S_i, which is modified through the mixing and separation processes, the weighting effects still maintain. As discussed in Sec.IV, the signal distortion is evaluated based on the distance from $H_{ji}S_i$. Therefore, when the signal source spectra are similar to each other, the above weighting can suppress signal distortion. If the signal sources exist in different frequency e.g. as is in music, their spectra are not similar and it can be expected the masking will cause signal distortion.

VI. Source Separation and Signal Distortion in FB-BSS

There are two cases, in which possible solutions for perfect separation exist, as shown below.

$$(1)\ C_{21}(z) = \frac{H_{21}(z)}{H_{11}(z)} \quad C_{12}(z) = \frac{H_{12}(z)}{H_{22}(z)} \quad (31)$$

$$(2)\ C_{21}(z) = \frac{H_{22}(z)}{H_{12}(z)} \quad C_{12}(z) = \frac{H_{11}(z)}{H_{21}(z)} \quad (32)$$

It is assumed that delay time of $H_{11}(z)$ and $H_{22}(z)$ are shorter than that of $H_{21}(z)$ and $H_{12}(z)$. This means that in Fig.2, the sensor of X_1 is located close to S_1, and the sensor of X_2 close to S_2. From this assumption, the solutions in case (1) become causal systems. On the other hand, the solutions in case (2) are noncausal.

When $C_{ij}(z)$ satisfy the separation conditions Eqs.(31), the output signals are given by

$$Y_1(z) = H_{11}(z)S_1(z) \quad (33)$$

$$Y_2(z) = H_{22}(z)S_2(z) \quad (34)$$

They are exactly the same as the criteria of the signal distortion discussed in Sec.IV. Therefore, the FB-BSS has a unique solution, which satisfies both the source separation and the signal distortion free simultaneously. Thus, in the FB-BSS, if complete signal separation is achieved, then signal distortion free is automatically satisfied.

VII. Simulation and Discussion

A. Simulation Conditions

The transfer function of the cross paths are related to the direct paths as follows:

$$H_{21}(z) = \alpha z^{-1} H_{11}(z) \quad (35)$$
$$H_{12}(z) = \alpha z^{-1} H_{22}(z) \quad (36)$$

Tow cases including $\alpha = 0.9$ (Mixture-1) and $\alpha = 0.5$ (Mixture-2) are taken into account. Mixture-1 is a difficult problem compared to Mixture-2. Speeches and colored signals, created by 2nd-order AR models, are used as sources. FFT size is 256 points in the frequency domain training. FIR filters with 256 taps are used in the FF-BSS, trained in the time domain and the FB-BSS. The initial guess of the separation block are $W_{11}(z) = W_{22}(z) = 1$ and $W_{ij}(z) = 0, i \neq j$, in the FF-BSS, and $C_{12}(z) = C_{21}(z) = 1$ in the FB-BSS.

Source separation is evaluated by the following two kinds of signal-to-interference ratios SIR_1 and SIR_2. $A_{ki}(z)$ is a transfer function from the i-th source to the k-th output. In this case, $S_1(z)$ and $S_2(z)$ are assumed to be separated in $Y_1(z)$ and $Y_2(z)$, respectively. However, it does not lose generarity.

$$\sigma_{s1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{11}(e^{j\omega})|^2 + |A_{22}(e^{j\omega})|^2)d\omega \quad (37)$$
$$\sigma_{i1} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{12}(e^{j\omega})|^2 + |A_{21}(e^{j\omega})|^2)d\omega \quad (38)$$
$$SIR_1 = 10\log_{10}\frac{\sigma_{s1}}{\sigma_{i1}} \quad (39)$$
$$\sigma_{s2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{11}(e^{j\omega})S_1(e^{j\omega})|^2 + |A_{22}(e^{j\omega})S_2(e^{j\omega})|^2)d\omega \quad (40)$$
$$\sigma_{i2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}(|A_{12}(e^{j\omega})S_2(e^{j\omega})|^2 + |A_{21}(e^{j\omega})S_1(e^{j\omega})|^2)d\omega \quad (41)$$
$$SIR_2 = 10\log_{10}\frac{\sigma_{s2}}{\sigma_{i2}} \quad (42)$$

B. Speech Signals

1) Source Separation: The learning curves of SIR_1 are shown in Figs.3, 4 and 5. Regarding convergence speed, the FF-BSSs are slightly faster than the FB-BSS. On the other hand, for separation performance, that is the value of SIR_1, the FB-BSS is superior to the others. Among the FF-BSSs, the frequency domain learning can provide better result.

2) Signal Distortion: The criteria for the signal distortion, that is the amplitude response of $H_{11}(z)S_1(z)$ and $H_{22}(z)S_2(z)$ are shown in Fig.6. The spectra of the output signals are shown in Figs.7, 8, 9 and 10. The mixing process is Mixture-2.

In the FF-BSS, trained in the time domain, the spectra are not similar to the criteria shown in Fig.6. Especially, the spectra in the high frequency band are amplified. Since, the FF-BSS has a degree of freedom, the output spectra can be changed in a way to make the output signals to be more statistically independent. On the other hand, as shown in Fig.8,

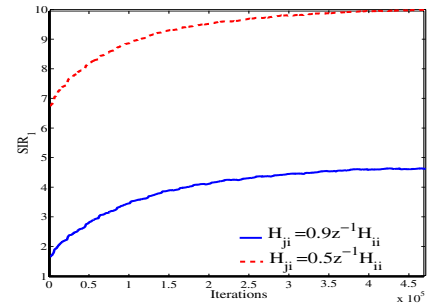

Fig. 3. Learning curve of SIR_1 in case of FF-BSS trained in time domain for speech signals.

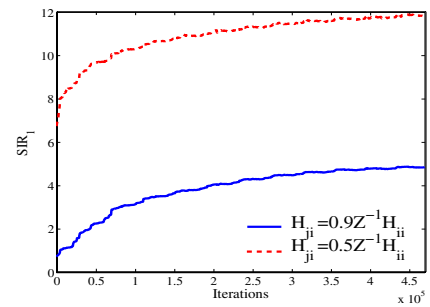

Fig. 4. Learning curve of SIR_1 in case of FF-BSS trained in frequency domain for speech signals.

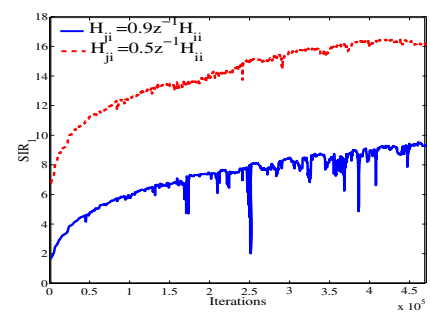

Fig. 5. Learning curve of SIR_1 in case of FB-BSS trained in time domain for speech signals.

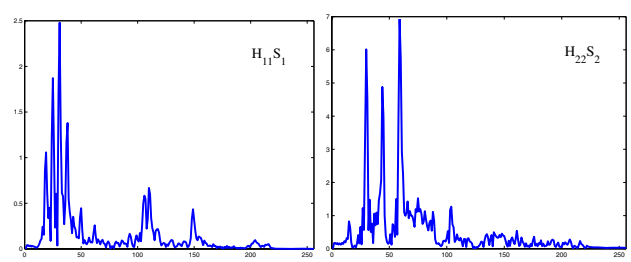

Fig. 6. Spectrum of $H_{11}(z)S_1(z)$ and $H_{22}(z)S_2(z)$ for speech signals.

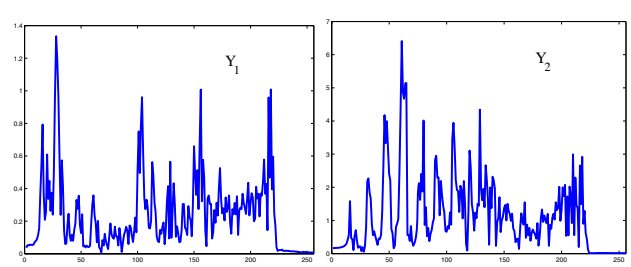

Fig. 7. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FF-BSS trained in time domain for speech signals

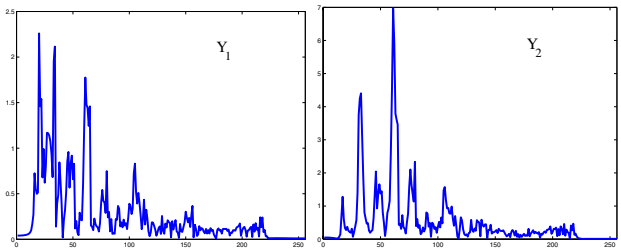

Fig. 8. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FF-BSS trained in time domain for speech signals.

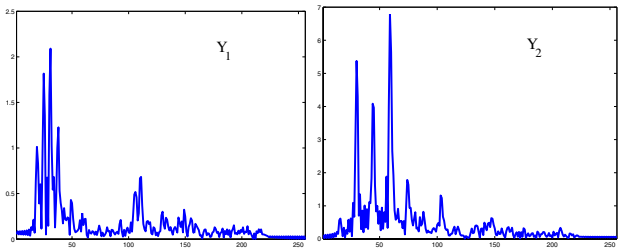

Fig. 9. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FF-BSS trained in frequency domain for speech signals.

the spectra of the FF-BSS, trained with the distortion free constraint, are drastically improved compared to the previous ones, and are similar to the criteria.

The FF-BSS, trained in the frequency domain, has a good result. Its output spectra are very similar to the criteria. As discussed in Sec.V-C, the separation block $\boldsymbol{W}(z)$ are trained using the output spectra as the mask. Furthermore, by using the identity matrix as the initial guess, the output spectra are similar to the observation in early stage of the learning process. This masking effect can suppress the signal distortion. Finally, the FB-BSS shown in Fig.10 also has a good result. As discussed in Sec.VI, the FB-BSS can satisfy the conditions of source separation and distortion free simultaneously. As sources are separated, signal distortion is also suppressed.

The other evaluation measures are summarized in TableI. Regarding SIR_2, the FF-BSS in frequency and the FB-BSS

TABLE I

COMPARISON OF FOUR KINDS OF BSSS FOR SPEECH SIGNALS. MIX-1AND 2 MEANS MIXTURE-1 AND 2, RESPECTIVELY. FF-BSS TIME(1), TIME(2) AND FREQ. ARE TRAINED FOLLOWING EQS.(3)-(4), EQS.(29)-(30) AND EQS.(5)-(6), RESPECTIVELY.

Methods	Mix	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-BSS	1	12.2	-0.49	-2.71	9.67	8.36
time(1)	2	19.3	-0.52	-2.82	10.4	9.22
FF-BSS	1	8.33	-6.20	-10.0	-11.1	-15.9
time(2)	2	15.6	-13.8	-16.8	-19.2	-22.8
FF-BSS	1	7.02	-3.15	-8.22	-9.20	-11.3
freq.	2	22.6	-16.5	-20.8	-20.3	-23.6
FB-BSS	1	14.1	-11.1	-14.3	-14.1	-16.3
	2	22.9	-23.3	-26.4	-22.8	-23.8

are the best. The signal distortion in the FF-BSS in time with distortion free constraint can be drastically improved compared to the FF-BSS in time without the constraint. Regarding SD_{1a},

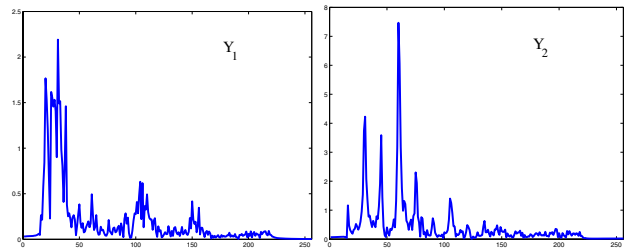

Fig. 10. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FB-BSS trained in time domain for speech signals.

which is the most strict evaluation, the FB-BSS is superior to the others. In SD_{1b}, which compare only amplitude responses, their differences become small, still the FB-BSS is superior to the others. In SD_{2a} and SD_{2b} evaluations, which compare only transfer functions, they are almost the same. Therefore, it can be concluded that regarding the signal distortion, the FB-BSS is the best according to any evaluation measures and any mixing models.

C. Colored Signals with Different Frequency Bands

As discussed in Sec.V-C, there is the weighting effect in the FF-BSS trained in the frequency domain. This weighting effect suppresses the signal distortion when the spectra of the sources are similar to each other, as a result, the spectra of the observed signals are also similar to those of the souces. However, it can be expected that the weighting is not effective for the sources, which have different envelop of the spectrum, and signal distortion will occur. In order to confirm this point, another example is shown here. Figs.11 and 12 show the spectra of the observed signals and the spectra of $H_{ii}(z)S_i(z)$, respectively. They are not similar to each other, because the frequency bands, where the spectra are dominant, are different.

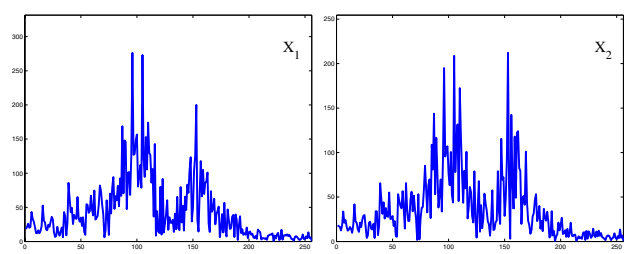

Fig. 11. Spectrum of observed signals for colored signals.

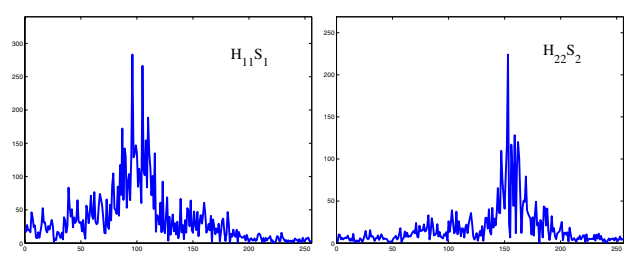

Fig. 12. Spectrum of $H_{11}(z)S_1(z)$ and $H_{22}(z)S_2(z)$ for colored signals.

The spectra of the output signals are shown in Figs. 13, 14, 15 and 16. The outputs of the FF-BSS trained in the frequency domain are not similar to the criteria, but are

similar to observed signals. This result supports our theoretical analysis. The other methods obtained similar results as in the simulations for speech signals.

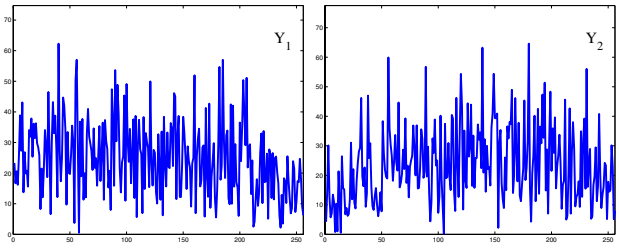

Fig. 13. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FF-BSS trained in time domain for colored signals.

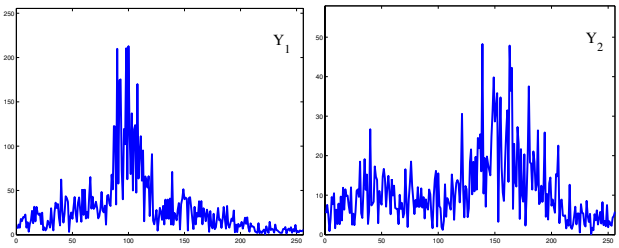

Fig. 14. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FF-BSS trained in time domain for colored signals.

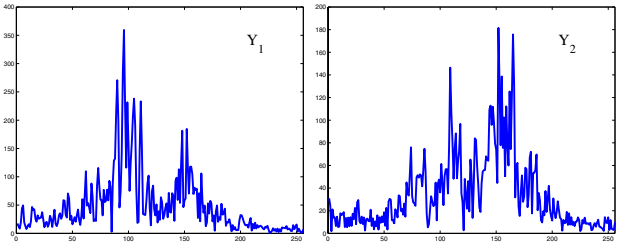

Fig. 15. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FF-BSS trained in frequency domain for colored signals.

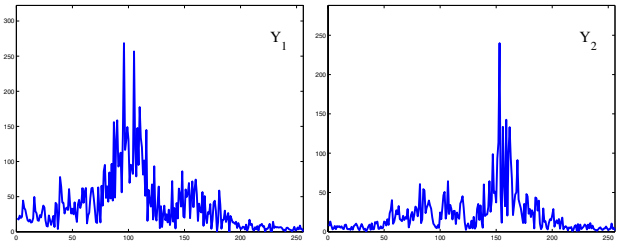

Fig. 16. Spectrum of output signals $y_1(n)$ and $y_2(n)$ in FB-BSS trained in time domain for colored signals.

VIII. Conclusions

In this paper, source separation and signal distortion in the FF-BSS and the FB-BSS have been analyzed. The new distortion free constraint has been proposed for the FF-BSS trained in the time domain. The FF-BSS, trained in the frequency domain, has the weighting effect, which can suppress signal distortion when the spectra of the sources are similar to each other. However, if the spectra of the sources differ from the others, the weighting is not effective, and signal distortion occurs. The FB-BSS, trained in the time domain, has a unique solution, which satisfies both the source separation and the distortion free conditions simultaneously. Simulation results support the theoretical analysis.

TABLE II

Comparison of four kinds of BSSs for colored signals. The mixing process is Mixture-1.

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-BSS time(1)	4.95	9.49	-0.08	-2.76	-0.69	-4.99
FF-BSS time(2)	7.07	10.2	-6.45	-9.58	-10.8	-13.6
FF-BSS freq.	2.62	5.12	-6.28	-8.74	-8.23	-10.1
FB-BSS	7.19	16.5	-12.4	-15.0	-10.4	-13.6

References

[1] C.Jutten and Jeanny Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," Signal Processing, 24, pp.1-10, 1991.

[2] E.Sorouchyari, "Blind separation of sources, Part III: Stability analysis," Signal Processing, 24, pp.21-29, 1991.

[3] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures," Signal Processing, vol.45, no.2, pp.209–229, March 1995.

[4] S.Amari, T.Chen and A.Cichocki, "Stability analysis of learning algorithms for blind source separation," Neural Networks, vol.10, no.8, pp.1345-1351, 1997.

[5] C.Simon, G.d' Urso, C.Vignat, Ph.Loubaton and C.Jutten, "On the convolutive mixture source separation by the decorrelation approach," IEEE Proc. ICASSP'98, Seattle, pp.IV-2109–2112, May 1998.

[6] S.Cruces and L.Castedo, "A Gauss-Newton methods for blind source separation of convolutive mixtures," IEEE Proc. ICASSP'98, Seattle, pp.IV2093–2096, May 1998.

[7] L.Parra and C.Spence, "Convolutive blind separation of nonstationary source," IEEE Trans. Speech Audio Processing, vol.8, pp.320-327, May 2000.

[8] H.Mathis and S.C.Douglas, "On optimal and universal nonlinearities for blind signal separation," IEEE Proc. ICASSP'01, MULT-P3.3, May 2001.

[9] S.Araki, S.Makino, T.Nishikawa and H.Saruwatari, "Fundamental limitation of frequency domain blind source separation for convolutive mixture of speech," IEEE Proc. ICASSP'01, Salt Lake City, MULT-P2.3, May 2001.

[10] I.Kopriva, Z.Devcic and H.Szu, "An adaptive short-time frequency domain algorithm for blind separation of nonstationary convolved mixtures," IEEE INNS Proc. IJCNN'01, pp.424-429, July 2001.

[11] K.Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by kurtosis for blind source separation," IEEE INNS, Proc. IJCNN'2002, Honolulu, Hawai, pp.1234-1239, May 2002.

[12] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm for convolutive blind source separation with transmission delay constraint," IEEE INNS, Proc. IJCNN'2002, Honolulu, Hawai, pp.1287-1292, May 2002.

[13] S.Araki, R.Mukai, S.Makino, T.Nishikawa, H.Saruwatari, "The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech," IEEE Trans. Speech and Audio Processing, vol.11, no.2, pp.109-116, March 2003.

[14] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm with adaptive exponential stepsize for blind source separation of convolutive mixtures with reverberations," IEEE INNS, Proc. IJCNN'2003 July 2003.

[15] K. Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output pdf for blind source separation," Proc. ICA2003, Nara, Japan, pp.427-432, April 2003.

[16] K.Nakayama, A.Hirano and Y.Dejima, "Analysis of signal separation and distortion in feedforward blind source separation for convolutive mixture," Proc. MWSCAS2004, Hiroshima, Japan, pp. III-207-III-210, July 2004.