# BLIND EQUALIZATION OF SWITCHING CHANNELS BY ICA AND LEARNING OF LEARNING RATE

*Howard Hua Yang* *Shun-ichi Amari*

Lab. for Information Representation, FRP, RIKEN
Hirosawa 2-1, Wako-shi, Saitama 351-01, JAPAN
Emails: hhy@koala.riken.go.jo, amari@zoo.riken.go.jp

## ABSTRACT

In the literature of blind equalization, algorithms developed for equalizing an SISO or SIMO channel fail sometimes when the channel condition is poor. We derive blind equalization algorithms from blind separation algorithms to equalize the SISO channel with fractionally sampling. The approach is also applied to equalize SIMO or MIMO channels. For switching channels, we use an updating rule to tune the learning rate of on-line algorithms automatically to follow the channel change. The idea is applicable to improve all blind equalization algorithms to equalize switching channels.

## 1. INTRODUCTION

Most of the blind equalization algorithms such as the super-exponential algorithms in [12] fail to equalize an ill-conditioned SISO channel.

When the fractional samples of the SISO channel output are available or the problem is to equalize an SIMO channel, we can first use the least-squares approach [13] to estimate the channel impulse responses, then use pseudo-inverse to recover the input sequence. This method can equalize some ill-conditioned channels. But it is not an on-line algorithm and may also fail sometimes.

The idea of the super-exponential method is to design a non-linear operation in the combined channel-equalizer domain which squeezes the combined system towards a $\delta$-function, then implement this operation in the equalizer domain. The method fails for some ill-conditioned SISO channels because the implementation is not perfect. One class of on-line algorithms for equalizing SIMO channels was proposed in [8] where the super-exponential method in [12] was perfectly implemented by the fractionally-spaced equalizer.

In this paper, we shall use a different approach to tackle the problem of equalizing SIMO channels. By reformulating the model for an SIMO channel output, we can use any blind separation algorithm based on independent component analysis (ICA) [7] for blind equalization. Especially, we can use those algorithms in [2, 3, 6] with equivariant property in order to equalize ill-conditioned channels. Therefore, from the blind separation algorithms, we obtain a class of on-line algorithms for equalizing SIMO channels or SISO channels with fractionally-sampled channel outputs.

The performance of the on-line blind equalization algorithms depends on the learning rate. The idea of learning of learning rate developed for neural networks [1, 4, 11] can be used to improve the blind equalization algorithms especially for switching channel or non-stationary channel.

The contents of this paper are organized in the following way. The model and the problem are described in Section 2. The blind equalization algorithms are given in Section 3: the algorithm based on ICA in Section 3.1 and the algorithm for learning rate in Section 3.2. To demonstrate the effectiveness of our algorithms, we show some simulation results in Section 4.

## 2. CHANNEL MODEL

Consider the following SISO channel with discrete time input and continuous time output:

$$x(t) = \sum_{kT\le t} s(k)h(t-kT) + n(t) \tag{1}$$

where $\{s(k)\}$ is an input sequence, T the symbol interval, $n(\cdot)$ the additive noise, and $h(\cdot)$ the channel impulse response function. Assume $h(t) = 0$ if $t \notin [0, LT]$. When the sampling rate is M times faster than the baud rate, the model (1) becomes an SIMO system:

$$x_m(k) = \sum_{l=0}^{L} h_m(l)s(k-l) + n_m(k), \tag{2}$$
$$k = 1, \cdots, N, m = 1, \cdots, M,$$

where $x_m(k)$ is the output of the m-th channel, N the data length, $n_m(k)$ a zero mean additive noise, and

$$h_m(l) = h(lT + \frac{m}{M}T), \quad l = 0, \cdots, L$$

are the impulse responses of the m-th channel.

Our problem is to equalize the SISO channel (1) with fractionally sampled outputs or the SIMO channel (3).

## 3. BLIND EQUALIZATION

### 3.1. Equalization via ICA

To recover the input only from the outputs of the SIMO system (3), we first reformulate this system as a linear mixture model

$$\boldsymbol{x}(k) = \boldsymbol{A}\boldsymbol{s}(k) + \boldsymbol{n}(k)$$

and then apply the following algorithm for de-mixing:

$$\boldsymbol{W}_{k+1} = (1+\eta)\boldsymbol{W}_k - \eta \mathbf{f}(\boldsymbol{y}_k)\boldsymbol{y}_k^T \boldsymbol{W}_k \quad (3)$$

where $\eta$ is a learning rate and

$$\mathbf{f}(\boldsymbol{y}) = (f(y_1), \cdots, f(y_n))^T$$

for some non-linear function $f(\cdot)$.

This algorithm has been derived in [2] by minimizing the mutual information of the outputs using the natural gradient descent method. The algorithms of this type can also be derived by other approaches such as info-max[5] and maximum likelihood. Different approaches may give different function forms for $f(\cdot)$ such as instantaneous functions in [3] and adaptive ones in [15].

Define $\underline{\mathbf{h}}_l = (h_1(l), \cdots, h_M(l))^T$, $l = 0, 1, \cdots, L$, and a $p \times (p+L)$ block matrix $\boldsymbol{H}_p$:

$$\begin{bmatrix} \underline{\mathbf{h}}_L & \underline{\mathbf{h}}_{L-1} & \cdots & \cdots & \underline{\mathbf{h}}_0 & 0 & \cdots & 0 \\ 0 & \underline{\mathbf{h}}_L & \underline{\mathbf{h}}_{L-1} & \cdots & \cdots & \underline{\mathbf{h}}_0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \underline{\mathbf{h}}_L & \underline{\mathbf{h}}_{L-1} & \cdots & \cdots & \underline{\mathbf{h}}_0 \end{bmatrix}.$$

The system (3) can be written as the following mixture model:

$$\boldsymbol{u}_k = \boldsymbol{H}_p \underline{\boldsymbol{s}}_k + \underline{\boldsymbol{n}}_k \quad (4)$$

where

$\boldsymbol{u}_k = [\boldsymbol{x}_k^T, \cdots, \boldsymbol{x}_{p+k-1}^T]^T$,
$\boldsymbol{x}_k = [x_1(k), \cdots, x_M(k)]^T$,
$\underline{\boldsymbol{s}}_k = [s(k-L), \cdots, s(p+k-1)]^T$,

and $\boldsymbol{n}_k$ is similarly defined as $\boldsymbol{u}_k$ but from the noise term in (3).

Note $\boldsymbol{H}_p$ is of $pM \times (p+L)$. When $p = L$ and $M = 2$, the $\boldsymbol{H}_p$ is a square matrix. This corresponds to the sampling interval $\frac{T}{2}$ which is popular in communication applications.

Assume $p \geq \frac{L}{M-1}$, $\boldsymbol{H}_p$ has a full column rank, and the input is an independent non-Gaussian sequence. This is the condition for using the blind separation algorithms for blind equalization. Let $H_m(z) = \sum_{k=1}^{L} h_m(k) z^{-k}$ be the transfer function of the m-th sub-channel. Under the condition $p(M-1) \geq L$, $\boldsymbol{H}_p$ has a full column rank if and only if the following zero condition [9] holds:

$h_m(0) \neq 0$ for some $1 \leq m \leq M$,
$h_m(L) \neq 0$ for some $1 \leq m \leq M$, and
$\{H_m(z)\}_0^M$ have no common zeros.

When $p > \frac{L}{M-1}$, the output dimension $pM$ of the system (4) is greater than its input dimension $p+L$. To reduce the dimensionality and preserve dimensionality of the signal space at the same time, we use a $(p+L) \times pM$ whitening matrix $\boldsymbol{V}_k$ to transform $\boldsymbol{u}_k$ to $\boldsymbol{v}_k$. $\boldsymbol{V}_k$ is updated by the following whitening algorithm:

$$\boldsymbol{V}_{k+1} = (1+\alpha)\boldsymbol{V}_k - \alpha \boldsymbol{v}_k \boldsymbol{v}_k^T \boldsymbol{V}_k \quad (5)$$

where $\alpha$ is a learning rate.

Considering the system (4) as a mixture model with $p+L$ independent sources, we apply the algorithm (3) to update the matrix $\boldsymbol{W}_k$ and recover the source vector by a linear transform:

$$\widehat{\underline{\boldsymbol{s}}}_k = \boldsymbol{W}_k \boldsymbol{V}_k \boldsymbol{u}_k.$$

Since both $\underline{\boldsymbol{s}}_k$ and $\boldsymbol{H}_p$ are unknown, we cannot obtain the exact inverse of $\boldsymbol{H}_p$. However, we can use any blind separation algorithm based on ICA to obtain

$$\boldsymbol{W}_\infty = \boldsymbol{D}\boldsymbol{P}(\boldsymbol{V}_\infty \boldsymbol{H}_p)^{-1},$$

a scaled and permuted inverse of $\boldsymbol{V}_\infty \boldsymbol{H}_p$, where $\boldsymbol{D}$ is a diagonal matrix with non-zero diagonal elements and $\boldsymbol{P}$ is a permutation matrix. With $\boldsymbol{W}_\infty$ and $\boldsymbol{V}_\infty$, we achieve the equalization and obtain a possibly delayed input in each dimension of the vector $\widehat{\underline{\boldsymbol{s}}}_k$. This algorithm is called blind separation for blind equalization (BSBE). When the channel is ill-conditioned, $\boldsymbol{H}_p$ is nearly singular. Taking advantage of the equivariant property of those algorithms in [1, 3, 6], we can still equalize ill-conditioned channels by BSBE.

Note the above approach can be applied to develop algorithms to equalize an MIMO channel [14].

### 3.2. Learning of learning rate

The performance of the algorithm (3) depends on the learning rate. The following equation for the learning rate is proposed in [4] for learning realizable dichotomies:

$$\eta_{k+1} = \eta_k + \alpha \eta_k (\beta f(\boldsymbol{x}_k, \boldsymbol{y}_k) - \eta_k) \quad (6)$$

where $\alpha, \beta > 0$ and $f(\boldsymbol{x}_k, \boldsymbol{y}_k)$ is an error function. When the error signal $f(\boldsymbol{x}_k, \boldsymbol{y}_k)$ is small, the dynamics of $\eta_k$ is close to the dynamical system

$$\frac{d\eta}{dt} = -\alpha \eta^2$$

which has a solution $\eta(t) = \frac{1}{\alpha t}$. This corresponds to the annealing rate $\eta_k = \frac{1}{\alpha k}$. However, when the error signal $f(\boldsymbol{x}_k, \boldsymbol{y}_k)$ increases, the learning rate $\eta_k$ also increases. The equalizer is automatically tuned by increase the learning rate to adapt itself to the changing environment.

This scheme is further developed in [1] and [11] for neural learning and applied in blind separation to extract sources from non-stationary mixtures caused by a switching mixing matrix [11]. We can also apply this scheme to all online blind equalization algorithms especially for equalizing a switching channel.

## 4. SIMULATION

Consider the blind equalization of FIR filters with the following impulse response functions respectively:

**(A)** $h = [\ 0.4\ 1\ -0.7\ 0.6\ 0.3\ -0.4\ 0.1]$,

**(B1)** $h_1 = [\ -0.0008,\ 0.03,\ -0.0036,\ -0.0497,\ 0.0562,\ 0.3193,\ 0.4868,\ 0.3599,\ 0.0779,\ -0.0708,\ -0.0260,\ 0.0378]$,

**(B2)** $\begin{bmatrix} h_2 \\ h_1 \end{bmatrix}$ where $h_1$ is the same as **(B1)** and

$$h_2 = [\,-0.0166,\ 0.0189,\ 0.0226,\ -0.0354,\ -0.0231,\ 0.1799,\ 0.4339,\ 0.4596,\ 0.2185,\ -0.0248,\ -0.0639,\ 0.0150].$$

**(B1)** and **(B2)** are the impulse responses of the same channel with the baud rate and twice the baud rate as the sampling rate respectively.

The channel (A) was used in [12] to demonstrate the effectiveness of the super-exponential methods by which a family of blind equalization algorithms were derived. Under certain conditions, these algorithms are super-exponentially fast. We tested one algorithm in this family, namely the algorithm (60)-(61) in [12] (to be called SEBE in this paper), to equalize the channels **(A)** and **(B1)**. It is shown in Figure 1 that the SEBE succeeds in equalizing the channels **(A)** but failed to equalize the channel **(B1)**. The reason is the poor channel condition of **(B1)**.

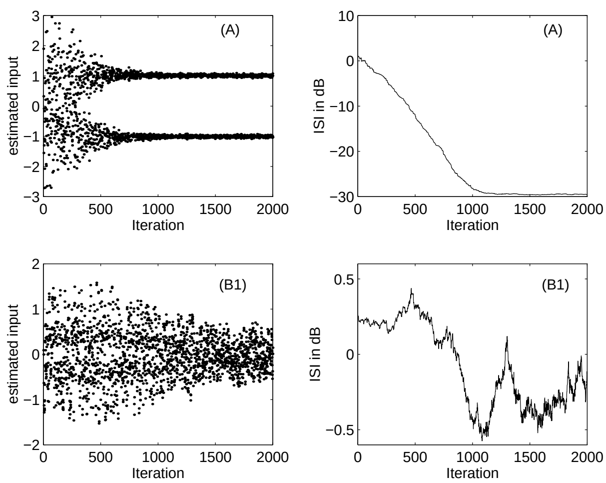


**Figure 1. Applying the SEBE for equalizing (A) and (B1)**

The zero distributions of $h$, $h_1$ and $h_2$ are plotted in Figure 2 to illustrate the channel conditions of **(A)** and **(B1)**. Needless to say that the channel **(B1)** and each sub-channel of the channel **(B2)** are ill-conditioned since both $h_1$ and $h_2$ have zeros on unit circle.

For the channel **(B2)**, although each sub-channel is ill-conditioned, we can first use the least-squares approach in [13] to estimate the channel impulse responses, and then use pseudo-inverse to estimate the input by the following: $\hat{\boldsymbol{s}}_k = (\widehat{\boldsymbol{H}}_p)^+ \boldsymbol{u}_k$. We call this algorithm the LSBE.

We test the both BSBE and LSBE for channel (B2). In the BSBE algorithm, we choose $M = 2$, $p = 11$, and $f(y) = y^3$. The outputs of BSBE and LSBE are shown in Figure 3.

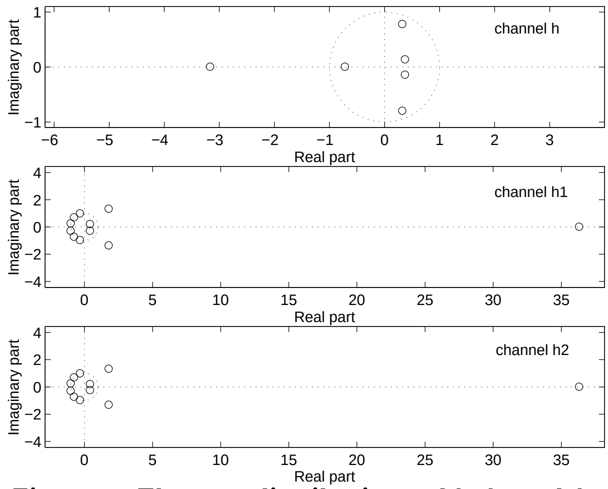


**Figure 2. The zero distributions of $h$, $h_1$ and $h_2$**

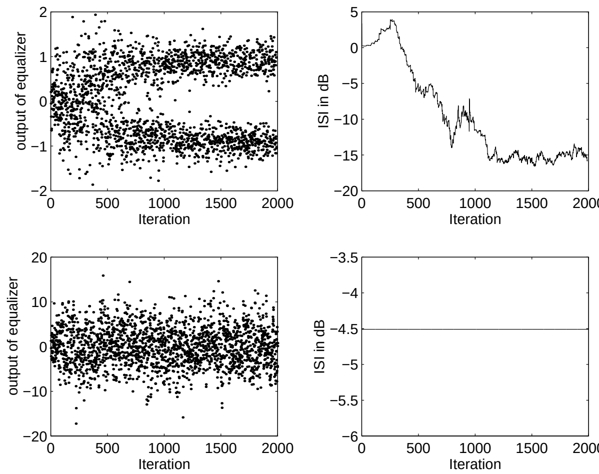


**Figure 3. Comparing BSBE to LSBE for equalizing (B2)**

The noise in the observation is a Gaussian noise with zero mean and variance $\sigma^2$. The input to the unknown channel is an independent binary sequence. When $\sigma = 0.0005$, after using 2000 symbols to equalize the channel we transmit 10000 symbols, count bit-error-rate (BER), and obtain BER=0.03% for BSBE and BER=23% for LSBE. This shows that BSBE equalizer performs better than LSBE equalizer at high SNR.

When the SNR is relatively low, the BSBE breaks down. This weakness is inherited from the noise problem in blind separation. Almost all blind separation algorithms are vulnerable to the noise added to the mixture. Nevertheless, since in many communication systems the SNR at the channel output is quite high and the dominant distortion is the ISI[10], the BSBE algorithm is still useful especially for equalizing ill-conditioned channels.

To demonstrate the effectiveness of the learning of learning rate, we consider the blind equalization of a switching channel in Figure 4. The two impulse response functions are $b_1 = [1,\ -5/2,\ -3/2]$ and $b_2 = [1,\ 5/3,\ -2/3]$. $\{c(k)\}$

is the equalizer and the channel observation $x(k)$ switches between two channels. We test the SEBE algorithm (60)-(61) in [12] combined with the learning of learning rate to equalize this channel. Instead of using the learning rate $\eta_k = \frac{1}{k}$ or a constant learning rate, we use the following scheme to update $\eta_k$:

$$\eta_{k+1} = \eta_k + \alpha \eta_k (\beta |y^2(k) - 1| - \eta_k)$$

It is shown in Figure 5 that the equalizer adapts itself to the changing environment by using the above learning rule for the learning rate.

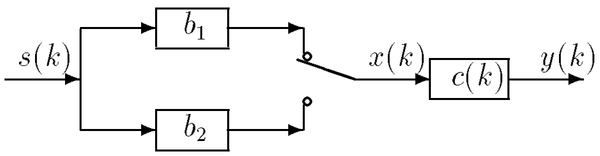


**Figure 4. A switching channel**

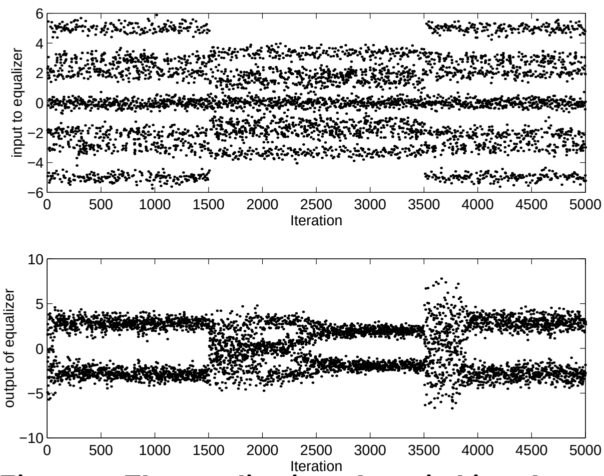


**Figure 5. The equalization of a switching channel**

## 5. CONCLUSIONS

We propose new on-line algorithms to equalize an SIMO channel or an SISO channel with fractionally sampling. The new algorithms have a merit to equalize an ill-conditioned channel because of the equivariant property inherited from the blind separation algorithms. This approach can be extended for blind equalization of MIMO channels. One simulation shows that the new algorithms perform better than the algorithm based on blind identification [13] and pseudo-inverse.

The idea of learning of learning rate can be applied to design algorithms for tuning blind equalization algorithms automatically in order to equalize a switching channel.

## REFERENCES


[1] S. Amari. Neural learning in structured parameter spaces – natural Riemannian gradient. In *Advances in Neural Information Processing Systems, 9, MIT Press: Cambridge, MA. (to appear)*, 1997.

[2] S. Amari, A. Cichocki, and H. H. Yang. A new learning algorithm for blind signal separation. In *Advances in Neural Information Processing Systems, 8, eds. David S. Touretzky, Michael C. Mozer and Michael E. Hasselmo, MIT Press: Cambridge, MA.*, pages 757–763, 1996.

[3] S. Amari, A. Cichocki, and H. H. Yang. Recurrent neural networks for blind separation of sources. In *Proceedings of NOLTA 1995*, volume I, pages 37–42, December 1995.

[4] N. Barkai, H. S. Seung, and H. Sompolinsky. Local and global convergence of on-line learning. *Physical Review Letters*, 75(7):1415–1418, August 1995.

[5] A. J. Bell and T. J. Sejnowski. An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, 7:1129–1159, 1995.

[6] J.-F. Cardoso and B. Laheld. Equivariant adaptive source separation. *IEEE Trans. on Signal Processing*, 43(12), December 1996.

[7] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36:287–314, 1994.

[8] Z. Ding. On convergence analysis of fractionally spaced adaptive blind equalizers. In *Proc. of 1996 ICASSP*, pages 2431–2434, April 1996.

[9] Y. Li and Z. Ding. Blind channel identification based on second-order cyclostationary statistics . In *Proc. of IEEE ICASSP'93*, pages IV–81–84, Minneapolis, MN, April 1993.

[10] Y. Li and Z. Ding. Convergence analysis of finite length blind adaptive equalizers. *IEEE Trans. on Signal Processing*, 43(9):2120–2129, September 1995.

[11] N. Murata, K. Muller, A. Ziehe, and S. Amari. Adaptive on-line learning in changing environments. In *Advances in Neural Information Processing Systems, 9, MIT Press: Cambridge, MA. (to appear)*, 1997.

[12] O. Shalvi and E. Weinstein. Super-exponential methods for blind deconvolution. *IEEE Trans. on Information Theory*, 39(2):504–519, March 1993.

[13] G. Xu, H. Liu, L. Tong, and T. Kailath. A least-squares approach to blind channel identification. *IEEE Trans. on Signal Processing*, 43(12):2982–2993, December 1995.

[14] H. H. Yang. On-line blind equalization via on-line blind separation. Technical Report BIP-96-0012, The Institute of Physical and Chemical Research (RIKEN), 1996.

[15] H. H. Yang and S. Amari. Two Gradient Descent Algorithms for Blind Signal Separation. In *Proceedings of ICANN96, The Lecture Notes in Computer Science Vol.1112, pp.287-292.* Springer-Verlag, 1996.