

# Blind Adaptation of Stable Discrete-Time IIR Filters in State-Space Form

Simone Fiori

**Abstract**—Blind deconvolution consists of extracting a source sequence and impulse response of a linear system from their convolution. In presence of system zeros close to the unit circle, which give rise to very long impulse responses, IIR adaptive structures are of use, whose adaptation should be carefully designed in order to guarantee stability. In this paper, we propose a blind-type discrete-time IIR adaptive filter structure realized in state-space form that, with a suitable parameterization of its coefficients, remains stable. The theory is first developed for a two-pole filter, whose numerical behavior is investigated via computer-based experiments. The proposed structure/adaptation theory is then extended to a multi-pole structure realized as a cascade of two-pole filters. Computer-based experiments are proposed and discussed, which aim at illustrating the behavior of the filter cascade on several cases of study. The numerical results obtained show the proposed filters remain stable during adaptation and provide satisfactory deconvolution results.

**Index Terms**—Blind system deconvolution; Bussgang-type deconvolution; Discrete-time adaptive filters; Inherently-stable IIR filters; State-space form.

## I. INTRODUCTION

**B**LIND system deconvolution (also known as blind signal restoration) is a challenging problem, which concerns designing the inverse of an unknown linear system driven by an unknown source signal while observing only its response signal [8], [9]. The user interest may be upon the forward system impulse response or on the source signal distorted by the forward linear system. Many engineering applications rely on blind deconvolution, such as channel equalization in telecommunication systems [20], blind image deblurring in digital image restoration [17], non-destructive material evaluation by ultrasonic inspection [22], remote sensing in geoscience [10], [23] and data storage and transmission [5], [18].

The majority of the available blind deconvolution techniques rely on finite-length impulse response (FIR) adaptive filters with a small number of taps [8], [9]. These structures are appropriate for mitigating the effects of systems whose zeros lie sufficiently far from the unit circle. In fact, the presence of zeros close to the unit circle would require very long inverse FIR filters and thus the use of infinite-length impulse response (IIR) filters would be beneficial in this case. However, IIR adaptive filters trained to have poles close to the unit circle may easily become unstable, therefore one of the research topics currently under investigation in adaptive IIR filtering is the challenging question of the design of *inherently-stable* filters. In the supervised adaptation field (as, e.g., in system identification) some convenient solutions have been developed in the recent past [7], [19]. A new solution, suited to on-line blind adaptation, is the subject of the present contribution.

This paper is organized as follows. Section II focuses on advancing the idea presented in the preliminary report [11] by introducing a different filter realization as well as new parameterization for the filter coefficients and by invoking a different adaptation criterion based on ‘Bussgang’ theory. A closed form expression for the total intersymbol interference (ISI), as a function of filter coefficients, and an

approximate closed-form expression for the Bussgang-type learning criterion, as a function of the ISI index, are provided. Numerical results illustrate the inversion of a two-zero system in a challenging case-study. Section III presents the extension of the second-order-section adaptation equations to the case of a cascade of an arbitrary number of two-pole filters. Such extension requires the development of a back-propagation-like rule for the simultaneous adaptation of the coefficients of all the filters in the cascade. The behavior of the proposed adaptation filter theory is investigated numerically via computer-based experiments. Finally, concluding thoughts are given in Section IV.

## II. FILTER STRUCTURE AND ADAPTATION THEORY FOR A SECOND-ORDER SECTION

We now illustrate the relevant equations related to the inherently-stable two-pole filtering structure, by briefly recalling the adopted blind-type adaptation theory. Additionally, we present an analytical study of the features of the devised algorithm in the two-pole case. The present section also illustrates the numerical behavior of the two-pole filter on some blind-deconvolution cases.

### A. Stable two-pole adaptive filter structure: Description and properties

The adaptive structure investigated in the present section is a second-order IIR section, having complex poles only, whose transfer function is given by:

$$W(z; n) = \frac{1}{(1 - \phi(n)z^{-1})(1 - \phi^*(n)z^{-1})}, \quad (1)$$

where  $\phi$  and  $\phi^*$  are the complex-conjugate poles of the adaptive IIR filter, which depend on the discrete-time index  $n \in \mathcal{Z}$ .

The classical stability conditions for static (time-invariant) linear filters are not sufficient – in general – to ensure the stability of a dynamical filter. The selection of an adequate time-domain implementation of the adaptive filter (1) is a key issue in order for the time-varying filtering structure to remain stable over time. A realization that enjoys this feature and exhibits favorable numerical properties [3], [19] is based on state-space equations. A two-pole-filter input-state-output representation is given by:

$$\begin{aligned} \begin{bmatrix} \xi_1(n+1) \\ \xi_2(n+1) \end{bmatrix} &= \rho(r(n)) \begin{bmatrix} \cos \theta(n) & -\sin \theta(n) \\ \sin \theta(n) & \cos \theta(n) \end{bmatrix} \times \\ &\quad \begin{bmatrix} \xi_1(n) \\ \xi_2(n) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} x(n), \\ u(n) &= \begin{bmatrix} c_1(n) & c_2(n) \end{bmatrix} \begin{bmatrix} \xi_1(n) \\ \xi_2(n) \end{bmatrix} + d(n)x(n). \end{aligned} \quad (2)$$

In the above equations,  $x(n) \in \mathcal{R}$  and  $u(n) \in \mathcal{R}$  denote the filter input and output sequence, respectively, while  $c_1(n) \in \mathcal{R}$ ,  $c_2(n) \in \mathcal{R}$  and  $d(n) \in \mathcal{R}$ ,  $\theta(n) \in [0, \pi)$  and  $r(n) \in \mathcal{R}$  are free parameters that actually describe the filter features, while  $\rho(\cdot) : \mathcal{R} \rightarrow (-1, +1)$ . Also, the quantities  $\xi_1(n) \in \mathcal{R}$  and  $\xi_2(n) \in \mathcal{R}$  denote the filter state

The author is with the Dipartimento di Elettronica, Intelligenza Artificiale e Telecomunicazioni (DEIT) of the Università Politecnica delle Marche, Via Breccie Bianche, I-60131 Ancona (Italy). E-mail address: fiori@deit.univpm.it.

variables. The filter input sequence is driven by the forward system output sequence, defined as:

$$x(n) = \sum_{k=0}^{L_h} h(k)s(n-k) + v(n), \quad (3)$$

where  $h(n)$  denotes the system impulse response of finite duration  $L_h$ ,  $s(n)$  denotes the source sequence and  $v(n)$  denotes an additive (measurement) noise.

The transfer function of the structure (2) is found to be:

$$\begin{aligned} W(z; n) &= \frac{d(n) + b_1(n)z^{-1} + b_2(n)z^{-2}}{1 - 2\rho(n)\cos\theta(n)z^{-1} + \rho^2(n)z^{-2}}, \\ b_1(n) &\stackrel{\text{def}}{=} c_1(n) - 2d(n)\rho(r(n))\cos\theta(n), \\ b_2(n) &\stackrel{\text{def}}{=} d(n)\rho^2(n) + c_2\rho(n)\sin\theta(n) - c_1(n)\rho(n)\cos\theta(n). \end{aligned} \quad (4)$$

By equating the transfer function (4) to the desired transfer function (1), it is readily seen:

$$\begin{aligned} c_1(n) &\stackrel{\text{def}}{=} 2\rho(r(n))\cos\theta(n), \\ c_2(n) &\stackrel{\text{def}}{=} \rho(r(n))\frac{\cos 2\theta(n)}{\sin\theta(n)}, \\ d(n) &\stackrel{\text{def}}{=} 1, \\ \phi(n) &= \rho(r(n))e^{i\theta(n)}, \end{aligned} \quad (5)$$

with  $i^2 = -1$ . It is worth remarking that, with the above definitions for the parameters  $c_1$ ,  $c_2$  and  $d$ , the only free filter parameters to adapt are  $r$  and  $\theta$ . About function  $\rho(r)$ , several choices are of course possible as, e.g.,  $\rho(r) \stackrel{\text{def}}{=} \tanh(r)$ .

### B. Bussgang-type adaptation theory

In order for a blind deconvolution problem to be consistent, the usual set of considered conditions is: 1) The random source sequence  $s(n)$  is zero-mean independent identically-distributed (IID). 2) The random source sequence is symmetric around zero and non-Gaussian. 3) The forward system has finite energy as does its inverse. Under these hypotheses, the blind deconvolution theory ensures that deconvolution is feasible, and the random source stream may be recovered up to arbitrary scaling and time-delay, namely after adaptation it should hold  $u(n) = Ks(n - \bar{n})$ , with  $K \in \mathcal{R}$  and  $\bar{n} \in \mathcal{Z}$  (the group-delay  $\bar{n}$  accounts for the total system-filter cascade lag) [15]. It is worth noting that in the basic model of concern here, no fractional lags are allowed.

In order to make the two-pole filter be able to perform blind system deconvolution, it is necessary to design a proper adaptation procedure. To this end, we rely on the definition of a criterion  $F = F(r, \theta)$  and of a Riemannian gradient-descent procedure to iteratively adapt the free filter parameters  $r$  and  $\theta$ . In the literature, known criteria for blind deconvolution are entropy-based [4], Bussgang [10], likelihood-based [6] and high-order statistics (HOS) [20], [21], [23]. In the present paper, we chose the Bussgang-type criterion, whose salient features are briefly recalled below. Bussgang criterion was deemed to be appropriate because it gives rise to fairly simple adaptation terms and to gain over author's previous experience in the field [10], [13].

During filter adaptation, the misadjustment of filter coefficients makes the filter output differ from the source signal. The misadjustment may be properly taken into account through the following filter output signal model:

$$u(n) = K(n)s(n - \delta(n)) + \mathcal{N}(n), \quad (6)$$

where  $K(n) \in \mathcal{R}$  and  $\delta(n) \in \mathcal{Z}$  denote the instantaneous scaling and lag, while  $\mathcal{N}(n)$  denotes *deconvolution noise*, whose suitable representation is a zero-mean, white, Gaussian random process, uncorrelated with the source signal [15]. Such representation is valid

under the assumption that the system-filter-cascade total impulse response is sufficiently long, which clearly holds for a IIR filter at the beginning of adaptation stage (please also see details in Appendix II). Using (6), an appropriate source signal estimator  $B(u) \approx Ks(n - \delta)$  can be designed according to Bayesian estimation theory. On the basis of the available estimator, an error criterion  $F$  may be constructed as:

$$2F \stackrel{\text{def}}{=} E[\mathcal{N}^2] = E[(u - B(u))^2]. \quad (7)$$

Over the last 10 years, several studies have shown how to choose a proper estimator on the basis of prior information on the deconvolution problem [10], [13], [15]. We hereafter denote  $F = E[\bar{F}(u)]$ , where the local criterion  $\bar{F}(u) \stackrel{\text{def}}{=} (u - \tanh(u))^2$  according, e.g., to [13].

In order to optimize the Bussgang criterion (7), several algorithms are available in the literature that include the class of gradient-based algorithms [9], [10] and fixed-point algorithms [12], [21]. On-line blind deconvolution algorithms usually rely on the former kind of adaptation procedures. In the present contribution, a particular Riemannian-gradient-type adaptation theory is exploited. The Riemannian gradient-based adaptation equations for the parameters  $r$  and  $\theta$  of the inherently-stable two-pole filter are:

$$\begin{bmatrix} r(n+1) \\ \theta(n+1) \end{bmatrix} = \begin{bmatrix} r(n) \\ \theta(n) \end{bmatrix} - \eta(n)\mathbf{G}^{-1}(r(n), \theta(n)) \begin{bmatrix} \frac{\partial F}{\partial r(n)} \\ \frac{\partial F}{\partial \theta(n)} \end{bmatrix}, \quad (8)$$

for  $n \geq 0$ . The sequence  $\eta(n) > 0$  denotes the adaptation stepsize and the positive-definite matrix-sequence  $\mathbf{G}(r, \theta)$  denotes the metric tensor that describes the geometry of the Riemannian parameter space. The selection of the metric tensor is discussed in subsection II-C along with the selection of appropriate stepsize schedules, which is tightly tied to it.

It is worth mentioning that the Riemannian-gradient-descent-type algorithm is closely related to both natural-gradient-based and Newton-type optimization theories. A discussion on the mutual relationships among these theories appeared recently in [14].

In order to compute the derivatives of the criterion function  $F$  with respect to filter coefficients, it is necessary to invoke an on-line algorithm for statistical expectation estimation updating when a new sample becomes available. A classical method is stochastic low-pass filtering, which for the pair of IIR filter parameters is:

$$\begin{aligned} \frac{\partial F}{\partial r(n)} &= \alpha \frac{\partial F}{\partial r(n-1)} + (1 - \alpha) \frac{\partial \bar{F}}{\partial r(n)}, \\ \frac{\partial F}{\partial \theta(n)} &= \alpha \frac{\partial F}{\partial \theta(n-1)} + (1 - \alpha) \frac{\partial \bar{F}}{\partial \theta(n)}, \end{aligned} \quad (9)$$

for  $n \geq 1$  with a pair of initial values that may be safely assumed equal to zero. The constant  $0 < \alpha < 1$  acts as a temporal smoothing coefficient. According to the state-space equations (2) and to the parameterization (5), the required derivatives with respect to parameter  $r$  read:

$$\begin{aligned} \frac{\partial \bar{F}}{\partial r(n)} &= \frac{\partial \bar{F}}{\partial u(n)} \frac{\partial u(n)}{\partial \rho(n)} \frac{d\rho(n)}{dr(n)}, \text{ where :} \\ \frac{\partial \bar{F}}{\partial \rho(n)} &= \frac{\partial c_1(n)}{\partial \rho(n)} \xi_1(n) + \frac{\partial c_2(n)}{\partial \rho(n)} \xi_2(n), \text{ with :} \\ \frac{\partial c_1(n)}{\partial \rho(n)} &= 2 \cos \theta(n), \quad \frac{\partial c_1(n)}{\partial \rho(n)} = \frac{\cos 2\theta(n)}{\sin \theta(n)}. \end{aligned} \quad (10)$$

Analogously, for the derivatives with respect to parameter  $\theta$ :

$$\begin{aligned} \frac{\partial \bar{F}}{\partial \theta(n)} &= \frac{\partial \bar{F}}{\partial u(n)} \frac{\partial u(n)}{\partial \theta(n)}, \text{ where :} \\ \frac{\partial \bar{F}}{\partial \theta(n)} &= \frac{\partial c_1(n)}{\partial \theta(n)} \xi_1(n) + \frac{\partial c_2(n)}{\partial \theta(n)} \xi_2(n), \text{ with :} \\ \frac{\partial c_1(n)}{\partial \theta(n)} &= -2\rho(r(n)) \sin \theta(n) \text{ and} \\ \frac{\partial c_2(n)}{\partial \theta(n)} &= -\frac{2 \sin 2\theta(n)}{\sin \theta(n)} - \frac{\cos 2\theta(n)}{\sin \theta(n) \tan \theta(n)}. \end{aligned} \quad (11)$$

The computation of the derivative  $\frac{d\rho}{dr}$  depends on selection of function  $\rho(\cdot)$  and is straightforward. The computation of the derivative  $\frac{\partial \bar{F}}{\partial u}$  is also straightforward and may be found in [10].

### C. Selection of metric tensor

It is known from differential geometry that the concept of gradient on a Riemannian manifold is defined up to a metric on the manifold or an inner product (see, e.g., the tutorial [14]). In the present case, the parameter manifold  $\mathcal{P}$  is formed by the pairs  $(r, \theta) \in \mathcal{R} \times [0, \pi)$ . The local metric structure of the parameter manifold  $\mathcal{P}$  is described by the metric tensor  $\mathbf{G}(r, \theta) \in \mathcal{R}^{2 \times 2}$ . If we denote by  $(\partial_r, \partial_\theta)$  a basis of differential operators of the tangent space  $T_{r, \theta} \mathcal{P}$  at  $(r, \theta)$  and by  $\langle \cdot, \cdot \rangle_{r, \theta} : T_{r, \theta} \mathcal{P} \times T_{r, \theta} \mathcal{P} \rightarrow \mathcal{R}$  the bilinear, symmetric inner product of two tangent vectors at  $(r, \theta)$ , the metric tensor is:

$$\mathbf{G}(r, \theta) = \begin{bmatrix} \langle \partial_r, \partial_r \rangle_{r, \theta} & \langle \partial_r, \partial_\theta \rangle_{r, \theta} \\ \langle \partial_\theta, \partial_r \rangle_{r, \theta} & \langle \partial_\theta, \partial_\theta \rangle_{r, \theta} \end{bmatrix}. \quad (12)$$

If the parameter manifold  $\mathcal{P}$  is endowed with standard Euclidean geometry, then the metric tensor is easily found to be  $\mathbf{G}_1(r, \theta) = \mathbf{I}_2$ . Another possible geometry is the one induced by the coordinate change  $\tilde{x} = \rho(r) \cos \theta$ ,  $\tilde{y} = \rho(r) \sin \theta$ , where  $(\tilde{x}, \tilde{y})$  belongs to the unit-disk  $\mathcal{D}$ . An Euclidean geometry for  $\mathcal{D}$ , namely  $\langle \partial_{\tilde{x}}, \partial_{\tilde{x}} \rangle_{\tilde{x}, \tilde{y}} = \langle \partial_{\tilde{y}}, \partial_{\tilde{y}} \rangle_{\tilde{x}, \tilde{y}} = 1$  and  $\langle \partial_{\tilde{x}}, \partial_{\tilde{y}} \rangle_{\tilde{x}, \tilde{y}} = 0$  induces a Riemannian geometry on  $\mathcal{P}$ . In fact, the above coordinate change yields:

$$\begin{aligned} \partial_r &= \rho'(r)(\cos \theta) \partial_{\tilde{x}} + \rho'(r)(\sin \theta) \partial_{\tilde{y}}, \\ \partial_\theta &= -\rho(r)(\sin \theta) \partial_{\tilde{x}} + \rho(r)(\cos \theta) \partial_{\tilde{y}}, \end{aligned} \quad (13)$$

The computation of the inner products of the basis operators  $\partial_r, \partial_\theta$  may be carried out by exploiting the bilinearity of the inner product and the orthonormality of the basis  $(\partial_{\tilde{x}}, \partial_{\tilde{y}})$  and results in the metric tensor:

$$\mathbf{G}_2(r, \theta) = \begin{bmatrix} \rho'^2(r) & 0 \\ 0 & \rho^2(r) \end{bmatrix}. \quad (14)$$

This metric tensor appears as a generalization of the one previously proposed by Amari and Douglas in [2]. It is worth noting that the inverse of metric tensor  $\mathbf{G}_2(r, \theta)$  is ill-conditioned and becomes singular for small as well as large values of  $r$  because  $\rho(r) \approx 0$  for  $r \approx 0$  and  $\rho'(r) \approx 0$  for large values of  $r$ . A close look at (9) reveals, for example, that for large values of the radius the rotated gradient is almost completely directed along the  $r$ -coordinate, thus large values of  $r$  are discouraged (in order to prevent instability). A known numerical way for overcoming this drawback is to replace the exact metric tensor with  $\mathbf{G}_2(r, \theta) + \epsilon^2 \mathbf{I}_2$ , where  $\epsilon \in \mathcal{R}$  is a small constant.

The selection of the most appropriate geometry in adaptive filtering depends on the application at hand. Two examples of Euclidean (metric tensor  $\mathbf{G}_1$ ) and induced (metric tensor  $\mathbf{G}_2$ ) gradient structures are shown in Figure 1, where solid lines represent level-contours of criterion  $F(r, \theta)$ , while arrows represent Riemannian gradient vector fields  $\mathbf{G}^{-1}(r, \theta) \frac{\partial F(r, \theta)}{\partial (r, \theta)}$ . The first case is an example of system with its zero laying far away from the unit circle. The second case is an example of system with its zero close to the unit circle. In both cases, the induced-gradient vector field looks more uniform than the Euclidean-gradient vector field: This facilitates the selection of an adaptation stepsize  $\eta(n)$  that is suitable far away from convergence as well as close to convergence. Also, the Euclidean gradient vector field looks bent toward the system zero only in its vicinity, while the induced vector field looks bent toward the cost function minimum almost everywhere in the explored region of parameter space.

### D. Theoretical performance analysis

In general, the assessment of the effectiveness of the proposed filtering structure may be effected through a deconvolution performance index. In the present paper, as a performance index we considered

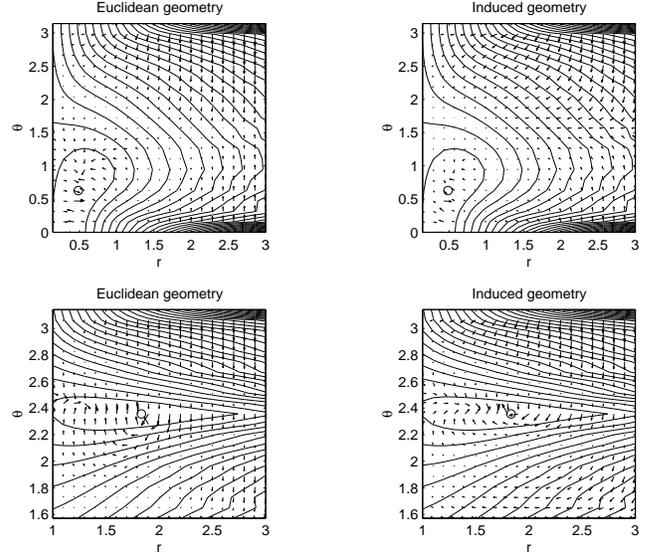


Fig. 1. Examples of Euclidean and induced gradient structures. Top panels: System zero far from the unit circle. Bottom panels: System zero close to the unit circle. (The vectors lengths are abs-log warped for graphical convenience.)

the inter-symbol interference (ISI) figure [20], defined as:

$$\text{ISI}(n) \stackrel{\text{def}}{=} \frac{\sum_k T_k^2(n)}{\max_k \{T_k^2(n)\}} - 1, \quad (15)$$

where  $T_k(n)$  denotes the global system-filter impulse response sequence at discrete-time  $n$ , namely, the convolution of the impulse responses of the system and filter (2). Also, for the theoretical performance analysis, it would be interesting to discuss the behavior of the Bussgang index  $F$  (7) as well as a signal restoration error, namely the mean squared error (MSE) between the filter output signal  $u(n)$  and the source signal  $s(n)$ .

Here, we present the exact closed-form expression of the ISI figure and MSE index, as well as a reasonable approximation of the Bussgang criterion, both as functions of the ISI figure. In particular, the latter expression allows judging the suitability of blind Bussgang index in the present context.

The working hypothesis for the analysis to be carried out are: 1) The forward filter to be cancelled is a two-tap FIR filter with transfer function  $H(z) \stackrel{\text{def}}{=} 1 - 2\rho_g \cos(\theta_g)z^{-1} + \rho_g^2 z^{-2}$ . The zeros of function  $H(z)$  are located at  $\rho_g e^{\pm i\theta_g}$ , therefore, by changing the value of  $\rho_g$  we can easily experiment with deconvolution of minimum-phase channels ( $\rho_g < 1$ ) as well as channels at the limit of phase-minimality ( $\rho_g$  close to 1). 2) The source signal  $s(n)$  is an IID sequence uniformly distributed within  $[-\sqrt{3}, +\sqrt{3}]$ .

The two-pole filter structure in the present section allows writing in closed form the ISI index as function of the filter pole  $\phi$  varying in the complex plane. The ISI figure (15) at time  $n$  reads:

$$\text{ISI}(n) = \frac{\rho_g^2}{|\phi(n)|^2} - 1 + 2\text{Re} \left\{ \frac{H(\phi(n))H(\phi^{-1}(n))}{(1-\phi^*(n)\phi^{-1}(n))(1-\phi^2(n))(1-|\phi(n)|^2)} \right\}. \quad (16)$$

The details of this calculation are give in Appendix I. It is straightforward to verify that the residual interference vanishes to zero for  $\phi = \rho_g e^{\pm i\theta_g}$ . Examples of the behavior of the ISI = ISI( $\phi$ ) function may be observed in Figure 2.

The Bussgang index at time  $n$  computes as:

$$F(n) = \int_{\mathcal{R}} p_u(u; n) \bar{F}(u) du, \quad (17)$$

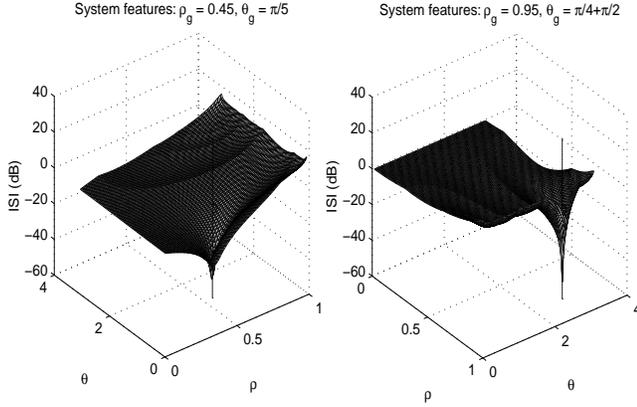


Fig. 2. ISI surface versus the parameters  $(\rho, \theta)$  for the experiments 1 and 2 on a two-zero/two-pole case.

where the function  $p_u(u; n)$  denotes the time-varying probability distribution of the filter output signal  $u(n)$ . Remarkably, with the indicated uniformly-distributed source signal, it is possible to find a reasonable approximation of the exact function  $p_u(u; n)$  in terms of the  $ISI(n)$  value only, namely:

$$p_u(u; n) \approx \frac{1}{4\sqrt{3}} \left[ \operatorname{erf} \left( \frac{u + \sqrt{3}}{\sqrt{2} ISI(n)} \right) - \operatorname{erf} \left( \frac{u - \sqrt{3}}{\sqrt{2} ISI(n)} \right) \right]. \quad (18)$$

The details of the above computation are given in Appendix II. For large values of ISI, which characterize the early stages of filter adaptation, each sample of filter's output signal computes as a linear combination of several source samples, therefore, by virtue of the central limit theorem of statistics, its instantaneous distribution should resemble a Gaussian one. Conversely, for small values of the ISI index, which should characterize the latest stages of filter adaptation, the distribution of the filter output values should closely resemble the distribution of the source signal, namely a uniform one. The adherence of the predicted distribution model (18) to the expected behavior may be verified on the left-hand panel of Figure 3.

On the basis of (18), the evaluation of the criterion function (17) may be obtained via numerical integration. Alternatively, for *large enough* values of the ISI index only, a sensible approximation of the output filter distribution is a zero-mean Gaussian of variance  $\sigma_u^2 = ISI$ , while an adequate approximation of the function  $\bar{F}(u)$  is:

$$\bar{F}(u) \approx \begin{cases} (u+1)^2 + \frac{1}{9}, & u < -1, \\ \frac{u^6}{9}, & |u| \leq 1, \\ (u-1)^2 + \frac{1}{9}, & u > +1, \end{cases} \quad (19)$$

which leads to the following approximate closed form of the Bussgang criterion as a function of the ISI index only:

$$F \approx \frac{5}{3} ISI^3 \operatorname{erf} \left( \frac{1}{\sqrt{2} ISI} \right) - \frac{5}{9} \sqrt{\frac{2 ISI}{\pi}} \exp \left( -\frac{1}{2 ISI} \right) (3 ISI^2 + ISI + 2) + \left( \frac{10}{9} + ISI \right) \left[ 1 - \frac{1}{\sqrt{\pi}} \Gamma \left( \frac{1}{2}, \frac{1}{2 ISI} \right) \right], \quad (20)$$

where symbol  $\Gamma(\cdot, \cdot)$  denotes the incomplete Gamma function [1]. The right-hand panel of Figure 3 suggests a relationship of proportionality between Bussgang criterion and the ISI index.

The mean squared error index is defined as:

$$\text{MSE} \stackrel{\text{def}}{=} E[(u - s)^2]. \quad (21)$$

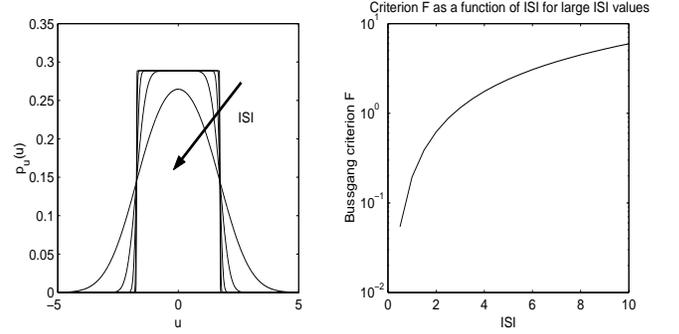


Fig. 3. Left: Approximated filter output distribution  $p_u(u)$  as a function of the ISI index, as given by (18). Right: Estimated relationship between the Bussgang criterion (17) and the ISI index (valid for large values of ISI).

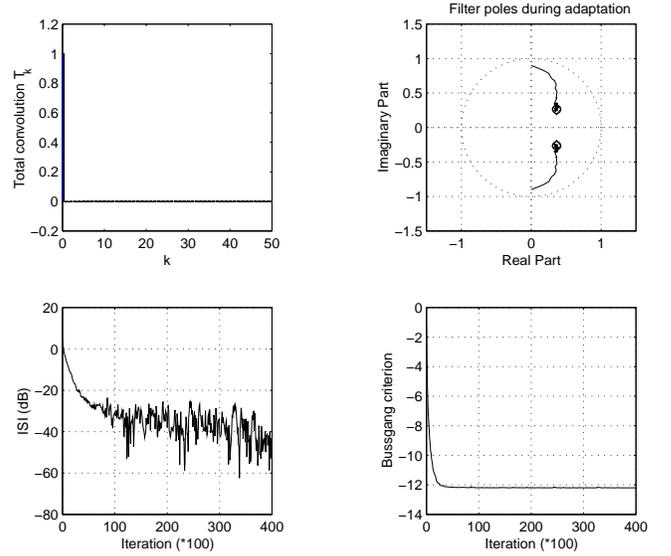


Fig. 4. Results of the experiment 1 on a two-zero/two-pole case. (The notation 'iteration (\*100)' means, e.g., that the plots encompass  $400 \times 100$  iterations.

It is worth noting, however, that in the noiseless case, the  $\text{MSE}(n)$  and  $ISI(n)$  indices are linearly related, namely  $\text{MSE}(n) = \sigma_s^2 ISI(n)$ , where  $\sigma_s$  denotes the variance of the source signal (which, in this case, equal 1). The details of the above computation are given in Appendix III.

#### E. Computer simulation results on a two-zero/two-pole case

The learning parameters for the following experiments were set to:  $\alpha = 0.8$ ,  $\eta(n) = 0.0005 \exp(-n/40000)$ ,  $r(0) = \operatorname{atanh}(0.9)$ ,  $\theta(0) = \frac{\pi}{2}$ .

The first experiment was performed with the following data:  $\rho_g = 0.45$ ,  $\theta_g = \frac{\pi}{5}$ . The results are shown in Figure 4. In the top-right panel, the system zeros are denoted by open circles, while the final filter pole locations are denoted by the diamonds, which are very close to the target zeros. In this experiment, after convergence the ISI index reaches  $-40$  dB. In particular, therefore, the mean squared error is of order  $10^{-4}$  that is very good in practice (see e.g. [6], [16] for a reference.) An important observation is that the ISI curve seems to oscillate unpleasantly around very low values: The reason of such oscillation may be recognized by observing the shape of the ISI surface versus the parameters  $(\rho, \theta)$  shown in the left-hand panel of Figure 2: The surface exhibits a very peaked shape around the optimal parameter-pair values, so that even a small deviation causes

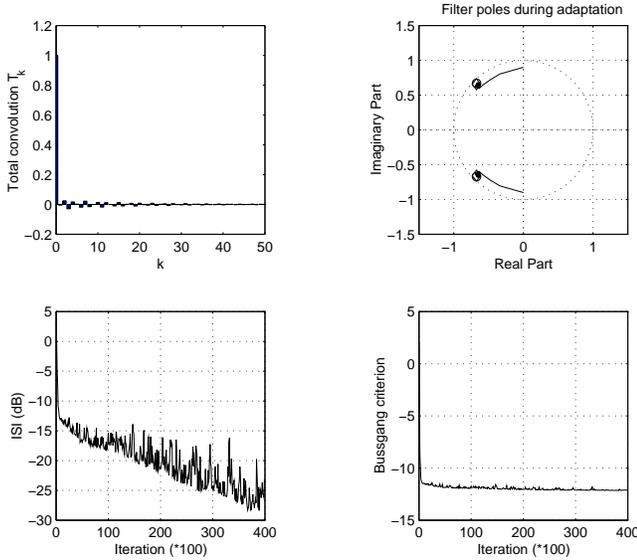


Fig. 5. Results of the experiment 2 on a two-zero/two-pole case.

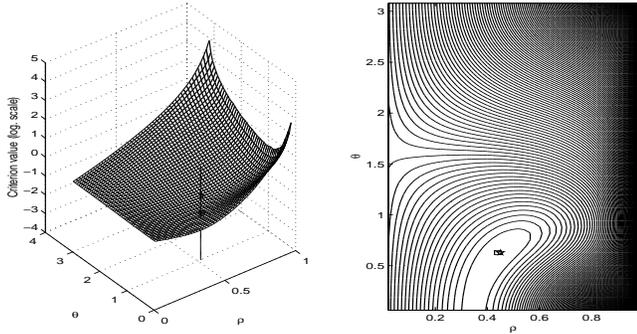


Fig. 6. Bussgang criterion surface versus the parameters  $(\rho, \theta)$  for the experiment 1 on a two-zero/two-pole case. (Star = True location of system zeros. Square = Estimated minimum of the criterion function.)

a large variation of the ISI value.

The second experiment was performed with the following data:  $\rho_g = 0.95$ ,  $\theta_g = \frac{\pi}{4} + \frac{\pi}{2}$ . The results are shown in Figure 5. The top-right panel shows that the final filter poles coincide to the target zeros with good faith. In this experiment, after convergence the ISI index is about  $-25$  dB that is good enough in practice.

It is worth evaluating numerically the shape of the Bussgang surface, as a function of the variables  $\rho$  and  $\theta$ , for the two cases of study just discussed. Figure 6 shows the criterion  $F$  surface for the data of the first experiment while Figure 7 refers to the second experiment. In both figures, the left-hand panel shows the tri-dimensional surface  $F = F(\rho, \theta)$  while the right-hand panel shows the level-curves representation of the same surface. Also, the star indicates the coordinates of the true zeros  $(\rho_g, \theta_g)$  while the square denotes the numerically-estimated minimum of the function  $F(\rho, \theta)$ . In both cases, the criterion functions appear to be convex around the optimal solutions and the estimated minima lie very close to the exact minima. This numerical analysis confirms the Bussgang criterion function as a good candidate for IIR-type blind deconvolution.

### III. EXTENSION TO A MULTI-POLE FILTER

On the basis of the adaptation theory developed for the second-order section in Section II, we now present an extension to a multi-pole filter, which is a cascade of second-order state-space filters. The

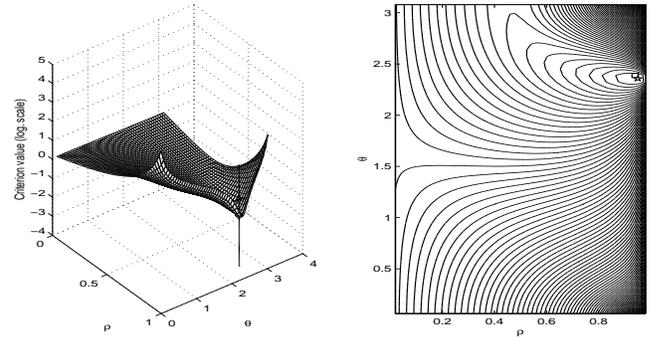


Fig. 7. Bussgang criterion surface versus the parameters  $(\rho, \theta)$  for the experiment 2 on a two-zero/two-pole case. (Star = True location of system zeros. Square = Estimated minimum of the criterion function.)

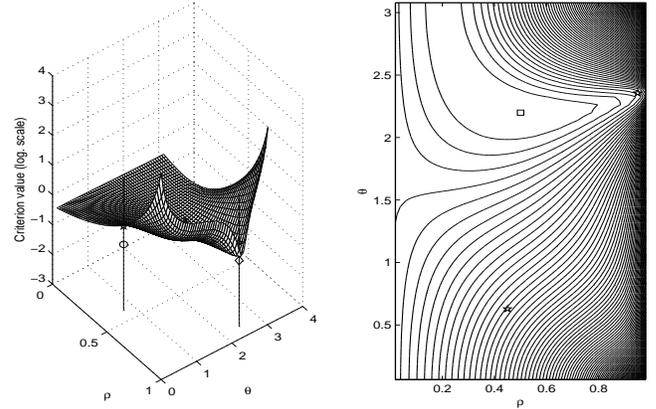


Fig. 8. Bussgang criterion surface versus the parameters  $(\rho, \theta)$  for a system with four zeros combined from experiment 1 and 2 of subsection II-E. (Stars = True locations of system zeros. Square = Estimated minimum of the criterion function.)

extension of the adaptation theory to the cascade of simple second-order sections will require the development of a sort of gradient-back-propagation algorithm for on-line adaptation of all second-order sections simultaneously.

#### A. On parallel versus sequential two-pole filters adaptation

Before developing a Bussgang-criterion gradient back-propagation algorithm for the multi-pole state-space adaptive filter, it is necessary to provide a rationale for simultaneous adaptation of the cascade.

A simpler choice would be to sequentially adapt individual second-order sections; that is, beginning to adapt a section when the preceding section in the cascade has completed adaptation. This procedure is unfeasible, though, because the combined presence of system zeros and second-order-filters poles warps the adaptation criterion surface as seen by next filters in the cascade. Such phenomenon is exemplified in the Figure 8, which shows the shape of the Bussgang surface – as a function of a section parameters  $\rho$  and  $\theta$  – in presence of a system having the four zeros combined from the two experiments of subsection II-E. In the right-hand panel of Figure 8, the crosses indicate the locations of the system zeros, while the square denote the minimum of the criterion  $F(\rho, \theta)$ . As it is readily seen, the presence of two zeros warps the criterion surface so that its minimum falls amidst the system zeros. This simple experiment leads to the conclusion that sequential simple-filter adaptation is not feasible in this context. In the case of parallel adaptation, the adaptation criterion function accounts for the joint

behavior of the single filters simultaneously and is therefore adequate to the simultaneous adaptation of their parameters.

### B. Stable multi-pole adaptive filter structure and adaptation

The number of second-order sections forming the cascade is denoted by  $L$  and the filter coefficients for the  $\ell^{\text{th}}$  section are now denoted as  $(r^{(\ell)}, \theta^{(\ell)})$ . The state-space equations for the  $\ell^{\text{th}}$  section read:

$$\begin{aligned} \begin{bmatrix} \xi_1^{(\ell)}(n+1) \\ \xi_2^{(\ell)}(n+1) \end{bmatrix} &= \rho(r^{(\ell)}(n)) \begin{bmatrix} \cos \theta^{(\ell)}(n) & -\sin \theta^{(\ell)}(n) \\ \sin \theta^{(\ell)}(n) & \cos \theta^{(\ell)}(n) \end{bmatrix} \\ &\quad \times \begin{bmatrix} \xi_1^{(\ell)}(n) \\ \xi_2^{(\ell)}(n) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} x^{(\ell)}(n), \\ u^{(\ell)}(n) &= \begin{bmatrix} 2\rho(r^{(\ell)}(n)) \cos \theta^{(\ell)}(n) & \rho(r^{(\ell)}(n)) \frac{\cos 2\theta^{(\ell)}(n)}{\sin \theta^{(\ell)}(n)} \end{bmatrix} \\ &\quad \times \begin{bmatrix} \xi_1^{(\ell)}(n) \\ \xi_2^{(\ell)}(n) \end{bmatrix} + x^{(\ell)}(n). \end{aligned} \quad (22)$$

The necessary cascade-connection equations read:

$$x^{(1)}(n) = x(n), \quad x^{(\ell)}(n) = u^{(\ell-1)}(n), \quad 2 \leq \ell \leq L. \quad (23)$$

The overall multi-pole filter output signal is clearly  $u^{(L)}(n)$ .

The Bussgang criterion is defined on the basis of the overall filter output signal, which should approximate the source signal after adaptation and is now written as  $F = E[\bar{F}(u^{(L)})]$ .

Each second-order filter parameter-pair is adapted according to the same gradient-based equation (8), where variables  $(\rho(n), \theta(n))$  are replaced by  $(\rho^{(\ell)}(n), \theta^{(\ell)}(n))$ .

For the required derivatives of the criterion function with respect to the parameters  $r^{(\ell)}$  and  $\theta^{(\ell)}$ , the results are the same of (9), (10) and (11), where variables  $(\rho(n), \theta(n), \xi_1(n), \xi_2(n))$  should be replaced by  $(\rho^{(\ell)}(n), \theta^{(\ell)}(n), \xi_1^{(\ell)}(n), \xi_2^{(\ell)}(n))$ .

In order to complete the calculation, it is necessary to compute the derivative  $\frac{\partial \bar{F}}{\partial u^{(\ell)}}$ . We distinguish between two cases. In the case  $\ell = L$ , the computation is straightforward, as the quantity  $\bar{F}$  is a function of the variable  $u^{(L)}$  only. In the other cases, namely for  $1 \leq \ell < L$ , we may use the following chain rules:

$$\begin{aligned} \frac{\partial \bar{F}}{\partial u^{(\ell)}} &= \frac{\partial \bar{F}}{\partial u^{(L)}} \frac{\partial u^{(L)}}{\partial u^{(\ell)}}, \\ \frac{\partial \bar{F}}{\partial u^{(\ell)}} &= \frac{\partial \bar{F}}{\partial u^{(L)}} \frac{\partial u^{(L)}}{\partial u^{(L-1)}} \frac{\partial u^{(L-1)}}{\partial u^{(\ell-1)}} \cdots \frac{\partial u^{(\ell+2)}}{\partial u^{(\ell+1)}} \frac{\partial u^{(\ell+1)}}{\partial u^{(\ell)}}. \end{aligned} \quad (24)$$

From the state-space equations (22) and the cascade-connection (23), it is straightforward to conclude that all the derivatives in the right-hand side of the last equation in (24) equal 1. Therefore, we conclude

$$\frac{\partial \bar{F}}{\partial u^{(\ell)}(n)} = \frac{\partial \bar{F}}{\partial u^{(L)}(n)}.$$

### C. Experiments with multi-pole IIR filters

In the following experiments, the filters coefficients were initialized as  $r^{(\ell)} = \text{atan}(0.9)$  and  $\theta^{(\ell)} = \frac{\pi \ell}{L+1}$  in order to cover the whole upper half-disk.

The first set of experiments aimed at verifying the ability of the multipole filter to deconvolve a system with four and six zeros located very close to the unit circle. For the four-zero system, the order of the filter was chosen equal to four ( $L = 2$ ) as well as six ( $L = 3$ ), in order to test for the robustness of the adaptation algorithm to model order mismatch. The results obtained in the four-zero/four-pole and the four-zero/six-pole case are depicted in Figures 9 and 10, respectively: In both situations, four adaptive filter poles match the system zeros; in the second case, the redundant pair of poles approaches the origin and give no contribution. The results obtained in the six-zero/six-pole case are depicted in Figure 11. All the presented results seem satisfactory as the value reached by the ISI index is below  $-20$  dB.

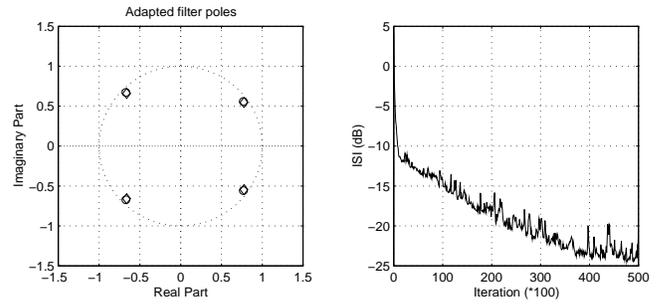


Fig. 9. Results of the experiment on a four-zero/four-pole case.

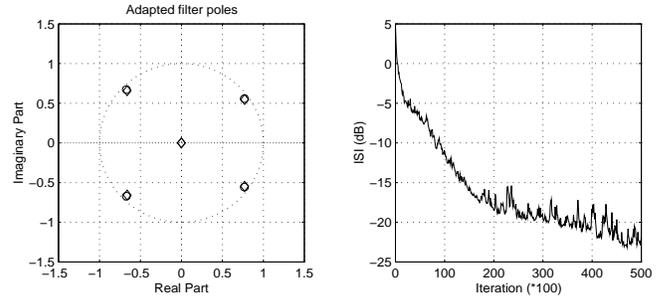


Fig. 10. Results of the experiment 2 on a four-zero/six-pole case.

An analysis of the behavior of an adaptive filter in presence of Gaussian noise  $v(n)$  added to system output, according to model (3), was also performed. The noise level is expressed via the signal-to-noise ratio (SNR) defined as:

$$\text{SNR} \stackrel{\text{def}}{=} \frac{E[s^2(n)]}{E[v^2(n)]}. \quad (25)$$

The average results (in terms of mean/variance of the ISI values recorded over 100 independent trials) are reported in the Table I. As long as the SNR level remains greater than 20 dB the result is completely satisfactory. With a SNR of less than 10 dB the performance seem to degrade progressively. The very low values of variances suggest the good stability of the adaptation process.

## IV. CONCLUSION

The aim of the present paper was to present an investigation on a possible adaptive inherently-stable IIR filtering theory for on-line blind signal deconvolution. The solution of choice was a cascade of two-pole filters, realized in state-space form with a special parameterization of the time-varying filter coefficients, in order to ensure stability. Single experiments as well as collective experiments on noiseless/noisy FIR forward systems, exhibiting zeros very close to the unit circle, illustrated the good behavior of the developed theory.

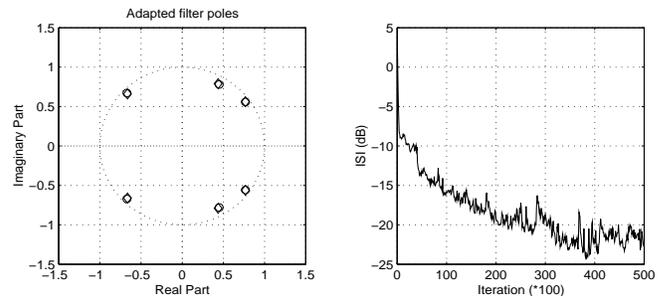


Fig. 11. Results of the experiment 2 on a six-zero/six-pole case.

TABLE I

AVERAGE RESIDUAL INTER-SYMBOL INTERFERENCE ISI, AS DEFINED IN (15), AND VARIANCE OVER 100 INDEPENDENT TRIALS VERSUS THE SIGNAL-TO-NOISE RATIO SNR, AS DEFINED IN (25), FOR A NOISY SIX-ZERO SYSTEM.

SNR (dB)	ISI AVE. (dB)	ISI VAR. (dB)
0	-1.488	0.175
10	-8.956	0.205
20	-19.943	0.334
30	-25.579	0.275
40	-25.694	0.033
50	-25.649	0.003

Extensions of the proposed method to mixed FIR/IIR filters with batch-type blind adaptation theories are currently under investigation, with the aim of growing the deconvolution capabilities of the proposed filter. Also, the extension of the SISO filter theory proposed in the present manuscript into a MIMO filter theory is under consideration.

## ACKNOWLEDGMENTS

The present version of the manuscript differs substantially from the first version submitted: The added material was conceived thanks to the interest into the treated topic displayed by the Associate Editor, Prof. Paul Fiore, as well as the anonymous Reviewers. I would like to gratefully thank them for the thorough comments and detailed suggestions.

Part of the present manuscript was prepared while the author was a visiting researcher at the Faculty of Information Technology, Mathematics and Electrical Engineering of Norwegian University of Science and Technology (Trondheim, Norway). The author wishes to gratefully thank Prof. Elena Celledoni for making this fruitful visit be possible and for pointing out reference [2].

The present version of the manuscript was prepared while the author was a short-term visitor of the Mathematical Neuroscience Laboratory of Brain Science Institute (BSI) at RIKEN (Wako-shi, Saitama, Japan) during July-September 2005: The author wishes to gratefully thank the BSI director, Prof. Shun-ichi Amari, and all the laboratory members for the warm and friendly hospitality.

## REFERENCES

- [1] M. ABRAMOWITZ AND C.A. STEGUN (Eds.), "Gamma (factorial) function" and "Incomplete Gamma function", Paragraphs 6.1 and 6.5 in Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9<sup>th</sup> printing, New York: Dover, pages 255 – 258 and 260 – 263, 1972
- [2] S.-I. AMARI AND S.C. DOUGLAS, *Why natural gradient?*, Proc. of Intl. Conf. on Acoustics, Speech and Sig. Proc. (ICASSP), Vol. 2, pp. 1213 – 1216, Seattle (WA, USA), May 12-15, 1998
- [3] C.W. BARNES AND A.T. FAM, *Minimum norm recursive digital filters that are free of overflow limit cycles*, IEEE Trans. on Signal Processing, Vol. 47, No. 9, pp. 2561 – 2567, Sept. 1999
- [4] A.J. BELL AND T.J. SEJNOWSKI, *An information maximisation approach to blind separation and blind deconvolution*, Neural Computation, Vol. 7, No. 6, pp. 1129 – 1159, 1995
- [5] A. BENVENISTE, M. GOURSAT, AND G. RUGET, *Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communication*, IEEE Trans. on Automatic Control, Vol. AC-25, No. 3, pp. 385 – 399, June 1980
- [6] A.M. BRONSTEIN, M.M. BRONSTEIN AND M. ZIBULEVSKY, *Relative optimization for blind deconvolution*, IEEE Trans. on Signal Processing, Vol. 53, No. 6, pp. 2018 – 2026, June 2005
- [7] P. CAMPOLUCCI AND F. PIAZZA, *Intrinsic stability-control method for recursive filters and neural networks*, IEEE Trans. on Circuits and Systems – Part II, Vol. 47, No. 8, pp. 797 – 802, Aug. 2000
- [8] C.-Y. CHI, C.-Y. CHEN, C.-Y. FENG AND C.-H. CHEN, *Batch processing algorithms for blind equalization using higher-order statistics*, IEEE Signal Processing Magazine, Vol. 20, No. 1, pp. 25 – 49, January 2003

- [9] A. CICHOCKI AND S.-I. AMARI, *Adaptive Blind Signal Processing*, Wiley 2003
- [10] S. FIORI, *A contribution to (neuromorphic) blind deconvolution by flexible approximated Bayesian estimation*, Signal Processing, Vol. 81, No. 10, pp. 2131 – 2153, Oct. 2001
- [11] S. FIORI, *Blind intrinsically-stable 2-pole IIR filtering*, Electronics Letters, Vol. 38, No. 23, pp. 1482 – 1483, Dec. 2002
- [12] S. FIORI, *Fast fixed-point neural blind deconvolution algorithm*, IEEE Trans. on Neural Networks, Vol. 15, No. 2, pp. 455 – 459, March 2004
- [13] S. FIORI, *Analysis of modified 'Bussgang' algorithms (MBA) for channel equalization*, IEEE Trans. on Circuits and Systems - Part I, Vol. 51, No. 8, pp. 1552 – 1560, Aug. 2004
- [14] S. FIORI, *Quasi-geodesic neural learning algorithms over the orthogonal group: A tutorial*, Journal of Machine Learning Research, Vol. 6, pp. 743 – 781, May 2005
- [15] S. HAYKIN, *Adaptive Filtering*, (Chapter 20: Blind deconvolution), Prentice-Hall, 1991
- [16] G.B. GIANNAKIS AND S. HALFORD, *Blind fractionally spaced equalization of noisy FIR channels: Direct and adaptive solutions*, IEEE Trans. on Signal Processing, Vol. 45, No. 9, pp. 2277 – 2292, Sept. 1997
- [17] D. KUNDUR AND D. HATZINAKOS, *Blind image deconvolution*, IEEE Signal Processing Magazine, Vol. 13, No. 3, pp. 43 – 64, May 1996
- [18] A.B. MARCHANT, *Optical Engineering: A Technical Overview*, Addison Wesley, 1996
- [19] P.A. REGALIA, *Adaptive IIR Filtering in Signal Processing and Control*, New York: Marcel Dekker, 1995
- [20] O. SHALVI AND E. WEINSTEIN, *New criteria for blind deconvolution of nonminimum phase systems (channels)*, IEEE Trans. on Information Theory, Vol. IT-36, No. 2, pp. 312 – 321, March 1990
- [21] O. SHALVI AND E. WEINSTEIN, *Super-exponential methods for blind deconvolution*, IEEE Trans. on Information Theory, Vol. 39, No. 2, pp. 504 – 519, 1993
- [22] S.K. SIN AND C.H. CHEN, *A comparison of deconvolution techniques for the ultrasonic non-destructive evaluation of materials*, IEEE Trans. on Image Processing, Vol. 1, No. 1, pp 3 – 10, Jan. 1992
- [23] R.A. WIGGINS, *Minimum entropy deconvolution*, Geoexploration, Vol. 16, pp. 21 – 35, 1978

## APPENDIX I

## APPENDIX A: CLOSED FORM EXPRESSION OF THE ISI INDEX FOR THE TWO-ZERO/TWO-POLE CASE

In order to derive the closed form expression of the ISI index for the two-zero/two-pole in the subsection II-E, let us first recall the relationship between the filter pole  $\phi$  and the overall causal impulse response  $T_k(n)$  at discrete-time  $n$ , which is  $T_k(n) = \mathcal{Z}^{-1}\{H(z)W(z;n)\}(k)$ , where  $\mathcal{Z}\{\cdot\}(\cdot)$  denotes the Z-transform operator. From this relationship it is possible to compute the ISI figure in closed form. In fact, by reversing it through the Fourier transform operator  $\mathcal{F}\{\cdot\}(\cdot)$ , we have  $\mathcal{F}\{T_k\}(\omega;n) = H(e^{i\omega})W(e^{i\omega};n)$ . Then, the Parseval theorem allows writing:

$$\sum_{k=0}^{+\infty} T_k^2(n) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} |H(e^{i\omega})W(e^{i\omega};n)|^2 d\omega. \quad (26)$$

This integral may be computed by the help of the residue theorem applied to the following equivalent line-integral over the unit-circle  $\partial\mathcal{D}$  in the complex plane:

$$\frac{1}{2\pi i} \oint_{\partial\mathcal{D}} H(z)H(z^{-1})W(z;n)W(z^{-1};n) \frac{dz}{z}. \quad (27)$$

The computation of the residues can be carried out by noting that the integrand, let us name it  $I(z;n)$ , has three single poles inside the unit-disk  $\mathcal{D}$ , namely in  $z = \phi$ ,  $z = \phi^*$  and  $z = 0$ . Residues are:

$$\begin{aligned} R_1 &\stackrel{\text{def}}{=} \text{Res}[I(z), 0] = \frac{|\phi_g|^2}{|\phi|^2}, \\ R_2 &\stackrel{\text{def}}{=} \text{Res}[I(z), \phi] = \frac{(1-\phi_g\phi^{-1})(1-\phi_g^*\phi^{-1})(1-\phi_g\phi)(1-\phi_g^*\phi)}{(1-\phi^*\phi^{-1})(1-|\phi|^2)(1-\phi^2)}, \\ R_3 &\stackrel{\text{def}}{=} \text{Res}[I(z), \phi^*] = R_2^*. \end{aligned} \quad (28)$$

Also, the initial-value theorem for the Z-transform may be used to compute  $\max_k T_k^2(n) = T_0^2(n)$ . Its application gives:

$$T_0(n) = \lim_{|z| \rightarrow +\infty} H(z)W(z; n) = 1. \quad (29)$$

It is worth noting that, in the two-zero two-pole case of concern here, the total impulse response has the maximum value at lag zero, therefore at perfect filter convergence the total delay of the system-filter cascade is null.

## APPENDIX II

### APPENDIX B: APPROXIMATION OF $p_u(u; n)$ AS A FUNCTION OF THE ISI FIGURE

The forward system to deconvolve is FIR, but the inverse adaptive filter is IIR, thus their cascade is IIR and the impulse response of the cascade system is  $T_k$ . The filter output signal is an infinitely-long linear combination of the samples of the source sequence, which are IID random variables, with combining coefficients  $T_k$ :

$$u(n) = \sum_{k=0}^{\infty} T_k(n)s(n-k). \quad (30)$$

Let us recall the definition of the first-kind characteristic function of a random variable  $x$ , that is  $\Phi_x(\omega) \stackrel{\text{def}}{=} \mathcal{F}\{p_x(x)\}(\omega) = E_x[e^{-i\omega x}]$ , where  $\omega \in \mathcal{R}$ . A useful property of the characteristic function is that  $\Phi_{c \cdot x}(\omega) = \Phi_x(c\omega)$  for every  $c \in \mathcal{R}$ .

According to the known combination rules of probability density functions for linear combination of random variables and to the duality theorem of Fourier transform, the filter output signal distribution  $p_u(u; n)$  may be written as:

$$\begin{aligned} p_u(u; n) &= \frac{1}{|\prod_{k=0}^{\infty} T_k(n)|} p_s\left(\frac{u}{T_0(n)}\right) \star p_s\left(\frac{u}{T_1(n)}\right) \star \\ & p_s\left(\frac{u}{T_2(n)}\right) \star p_s\left(\frac{u}{T_3(n)}\right) \star \dots \\ &= \mathcal{F}^{-1}\{\Phi_s(\omega) \prod_{k=1}^{\infty} \Phi_s(T_k(n)\omega)\}(u), \end{aligned} \quad (31)$$

where symbol ‘ $\star$ ’ denotes convolution. In order to develop the above calculations, we hypothesized that, at any time  $n$ , the global system-filter cascade impulse response is such that  $T_k(n) \neq 0$  for all  $k$ . As it is readily recognized, if some  $T_k = 0$ , the relative terms in the above equation disappear. It is worth noting that the statistics of the random signal  $u(n)$  are time-varying and therefore the probability density function depends on time.

It is possible to find a suitably approximated expression for the probability density  $p_u(u; n)$  for a uniformly-distributed source signal with variance  $\sigma_s^2 = 1$ . In this case, the characteristic function of the source signal has expression  $\Phi_s(\omega) = \frac{\sin(\sqrt{3}\omega)}{\sqrt{3}\omega}$ . A reasonable approximation of the above characteristic function is the Gaussian kernel  $\exp\left(-\frac{\omega^2}{2}\right)$ , therefore the infinite product in the expression (31) may be approximated by:

$$\prod_{k=1}^{\infty} \Phi_s(T_k\omega) \approx \exp\left(-\frac{\omega^2}{2} \sum_{k=1}^{\infty} T_k^2\right) = \exp\left(-\frac{\omega^2}{2} \text{ISI}\right). \quad (32)$$

The inverse transform of a Gaussian kernel is again a Gaussian kernel, therefore:

$$p_u(u; n) \approx p_s(u) \star \left[ \frac{1}{\sqrt{2\pi \text{ISI}(n)}} \exp\left(-\frac{u^2}{2 \text{ISI}(n)}\right) \right]. \quad (33)$$

The above convolution may be computed easily and leads to (18).

## APPENDIX III

### APPENDIX C: RELATIONSHIP BETWEEN ISI AND MSE INDICES

In the noiseless case, i.e., when  $v(n) = 0$  in the system output model (3), the difference  $u(n) - s(n)$  in the definition of MSE (21) may be written as:

$$u(n) - s(n) = \sum_{k=1}^{\infty} T_k(n)s(n-k), \quad (34)$$

where  $T_k(n)$  denotes again the global system-filter cascade impulse response. On the basis of this relationship, the MSE index recasts as:

$$\text{MSE}(n) = \sum_{k=1}^{\infty} \sum_{h=1}^{\infty} T_k(n)T_h(n)E[s(n-k)s(n-h)]. \quad (35)$$

As the source sequence is, by hypothesis, IID and zero-mean, its autocorrelation  $E[s(n-k)s(n-h)]$  is zero for  $k \neq h$  and equals  $\sigma_s^2 \stackrel{\text{def}}{=} E[s^2]$  for  $k = h$ .

In the noisy case, similar calculations show that the discussed relationship becomes:

$$\text{MSE}(n) = \sigma_s^2 \text{ISI}(n) + \sigma_v^2(n), \quad (36)$$

where  $\sigma_v^2(n)$  denotes the filtered additive system noise sequence  $\tilde{v}(n) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} T_k(n)v(n-k)$ .

PLACE  
PHOTO  
HERE

**Simone Fiori** was born in Rimini (Italy) in June 1971. He received the Italian Laurea (Dr.Eng.) *cum laude* in electronics engineering in July 1996 from the University of Ancona (Italy), and the Ph.D. degree in electrical engineering (circuit theory) in March 2000 from the University of Bologna (Italy). In November 2000 he has been appointed assistant professor of circuit theory at the University of Perugia (Italy). His research interests include unsupervised learning theory for artificial neural networks, linear and non-linear adaptive discrete-time filter theory, vision and image processing by neural networks, continuous-time and discrete-time circuits for stochastic information processing. He is author of more than 100 refereed journal and conference papers on these topics. Dr. Fiori has been the recipient of the 2001 ‘‘E.R. Caianiello Award’’ for the best Ph.D. dissertation in the artificial neural network field and is currently serving as Associate Editor of Neurocomputing journal.