A Robust Orthogonal Adaptive Approach to SISO Deconvolution

P. D. Doñate

Departamento de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur, Av. Alem 1253,
Bahía Blanca 8000, Argentina
Email: pdonate@criba.edu.ar

C. Muravchik

Facultad de Ingeniería, Universidad Nacional de La Plata; Comisión de Investigaciones Científicas de la Provincia
de Buenos Aires (CIC), La Plata 1900, Buenos Aires, Argentina
Email: chmuravchik@ieee.org

J. E. Cousseau

Departamento de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur, Av. Alem 1253,
Bahía Blanca 8000, Argentina

Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Cdad. de Buenos Aires 1033, Argentina
Email: iecousse@criba.edu.ar

This paper formulates in a common framework some results from the fields of robust filtering, function approximation with orthogonal basis, and adaptive filtering, and applies them for the design of a general deconvolution processor for SISO systems. The processor is designed to be robust to small parametric uncertainties in the system model, with a partially adaptive orthogonal structure. A simple gradient type of adaptive algorithm is applied to update the coefficients that linearly combine the fixed robust basis functions used to represent the deconvolver. The advantages of the design are inherited from the mentioned fields: low sensitivity to parameter uncertainty in the system model, good numerical and structural behaviour, and the capability of tracking changes in the systems dynamics. The linear equalization of a simple ADSL channel model is presented as an example including comparisons between the optimal nominal, adaptive FIR, and the proposed design.

Keywords and phrases: SISO deconvolution, robust filters, orthogonal basis, adaptive filters.

1. INTRODUCTION

Deconvolution filters have a wide range of applications in communications, control and signal processing. Among other roles, they are used to reduce the distortion and additive noise that contaminate a signal propagating through some channel. When noise levels are negligible and the transmission part of the system is minimum phase and perfectly known, these filters are obtained as the inverse of the original system. However, if the system is nonminimum phase and noise is also present, a realizable deconvolution filter, that is, a filter that is stable and causal, and uses a finite smoothing lag, cannot achieve perfect signal reconstruction. In such case, the design procedure focuses on minimizing some performance index and thus, different optimal deconvolution filters are possible according to the objective functions used.

Another source of difficulty are the uncertainties in the model of the transmission path or in the noise spectrum.

This phenomenon is related to modeling errors, noise in the data used for identification, the random nature of the noise description, and other physical causes as time variations, changing environments, component aging, and drift. The deconvolution filter has to be capable of tracking these changes or exhibit a robust behavior assuring a good performance to the extent of these variations.

Different ways of dealing with these problems are available in the literature. If there is very little knowledge about the system, then blind or blind adaptive techniques have to be used [1, 2]. These methods rely in random models and make use of the statistical theory for signal separation. When the system can be described by uncertain parametric models, robust approaches are available. In [3] a mean square error (MSE) is averaged with respect to model errors and noise. Probabilistic descriptions of the models uncertainties are used and the problem is formulated and solved by means of a polynomial approach. These results are further extended

to nonlinear equalization applications in [4] and presented as a general polynomial equations framework for nominal and robust multivariable linear filtering in [5]. The problem of nonlinear equalization is also addressed in [6] where a design method for decision feedback equalizers (DFE) to be applied in transmission systems with small parameter perturbations is presented. A simple probabilistic structure for channel and noise models is used and then a closed-form result in the frequency domain using calculus of variation and spectral factorization is derived. This same methodology is used in [7] to solve the problem of linear deconvolution.

All these approaches yield time-invariant, that is, fixed, recursive structures for the optimal filters. However, in applications where the environment may suffer larger changes, the filters will also require some degree of adaptation. Time-varying or adaptive deconvolution filters are the common solution to this problem, increasing complexity, computational load, and cost. This type of solutions usually involves the use of transversal or finite impulse response (FIR) adaptive filters as an approximation to naturally recursive systems.

The contribution of this paper is the formulation of a comprehensible framework that concentrates some of the results given in the fields of robust filter design, function approximation by orthogonal bases, and adaptive filtering. The aim is the design of a general deconvolution processor, robust to parametric uncertainties in the system model and with a partially adaptive recursive orthonormal structure.

Robustness focuses on assuring a reasonable performance over the range of "practical restricted complexity parameterized system models," a set of rational functions identified from a finite noisy data record, and gaining properties similar to the designs of [3] or [7]. The recursive orthogonal structure has a twofold function. First, it approximates recursive systems naturally, requiring less parameters than FIR approximations. Second, it gives the design the classical advantages of orthonormal bases, that is, modularity, good numerical conditioning, and simplified performance analysis [8] along with other practical properties [9]. Adaptation is intended to extend the range of applicability of the design. Simple strategies can be used exploiting the orthogonal structure and updating only the coefficients that combine the basis functions. Because of its recursive nature, the performance can be close to full adaptivity with a lower computational load than that required by long FIR adaptive filters [10, 11, 12, 13].

The design procedure is based on the optimization of a performance index that contemplates both the system model uncertainties and the usual quadratic error. The formulation is similar to that introduced in [6] for the nonlinear DFE and close to the development presented in [7]. The minimization follows the classical approach in the frequency domain and uses variational concepts. The results are presented in a theorem that establishes the optimum set of parameters of the robust orthogonal deconvolution processor.

The orthonormal structure is provided by time-invariant basis functions that have a simple construction [14] and allow the inclusion of different modes (poles). Adaptation is provided by a simple "gradient" updating algorithm. This algorithm updates the coefficients that linearly combine the basis functions. Some preliminary results in relation with this type of formulation were presented in [15] for a simplified deconvolution setup and in [16] for the application of echo cancellation. In this case, a fixed orthogonal basis (Laguerre) with a transversal filter type of adaptive structure was used for updating the coefficients.

The paper is organized as follows. Section 2 introduces some notation, general considerations, and the basis functions. The main results are developed in Section 3. Section 4 considers the coefficients updating algorithm. Section 5 presents an example where the proposed design strategy is used to derive an equalizer for a simple ADSL communication channel model. Comparisons of performance are made in terms of the MSE that different designs can theoretically achieve. Finally, in Section 6, some conclusions are drawn.

2. THE SISO DECONVOLUTION PROBLEM

2.1. Notation and general description

Most common SISO deconvolution or inverse filtering problems are described by the simple scheme illustrated in Figure 1 where the signals involved are modeled:

$$x(k) = H(q^{-1}, \boldsymbol{\alpha})a(k) + v(k),$$
$$v(k) = D(q^{-1}, \boldsymbol{\beta})n(k),$$
$$a(k) = W(q^{-1})d(k), \tag{1}$$
$$s(k) = \lfloor T(q^{-1})p(k) + a(k)\rfloor q^{-l},$$

$$\sigma_d^2 = E\{d^2(k)\}, \qquad \sigma_n^2 = E\{n^2(k)\},$$
$$\sigma_p^2 = E\{p^2(k)\}. \tag{2}$$

$H(q^{-1}, \boldsymbol{\alpha})$ and $D(q^{-1}, \boldsymbol{\beta})$ are linear time-invariant filters that form the *system*. They are functions of q^{-1}, the unitary delay operator, that is, $q^{-1}f(k) = f(k-1)$. These filters are not known exactly in the sense that they also depend on unknown real parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. We will use the simplified notation H and D when this dependence does not need to be put explicitly into evidence or, for example, $H(\boldsymbol{\alpha})$ and $D(\boldsymbol{\beta})$ when the time information is not central in an argument. The same considerations apply when working in the transform domain with $\mathcal{Z}\{f(k-1)\} = z^{-1}F(z)$. For example, the following representations of $H(z^{-1}, \boldsymbol{\alpha})$ are equivalent when used in the right context: H, $H(z^{-1})$, and $H(\boldsymbol{\alpha})$.

The input shaping filters $W(q^{-1})$ and $T(q^{-1})$ are perfectly known invertible linear filters that model the stochastic sequences $a(k)$ and $c(k)$. The signals $d(k)$, $p(k)$, and $n(k)$ are mutually independent, zero mean white stochastic sequences with variance σ_d^2, σ_p^2, and σ_n^2, respectively. The symbol $*$ is used to denote complex conjugation on $|z| = 1$ and transposition, so that if $\mathbf{G}(z^{-1}, \boldsymbol{\alpha})$ is a matrix of rational functions, then $\mathbf{G}^* = \mathbf{G}^*(z^{-1}, \boldsymbol{\alpha}) = \mathbf{G}^T(z, \boldsymbol{\alpha})$. The analytic part of H outside (resp., inside) the unit circle is denoted by $\{H\}_+$ (resp., $\{H\}_-$). The degree of a polynomial is indicated as $\mathcal{O}(\cdot, \cdot)$, where the arguments stand for negative and positive powers of q (or z) in that order. If only one argument is used,
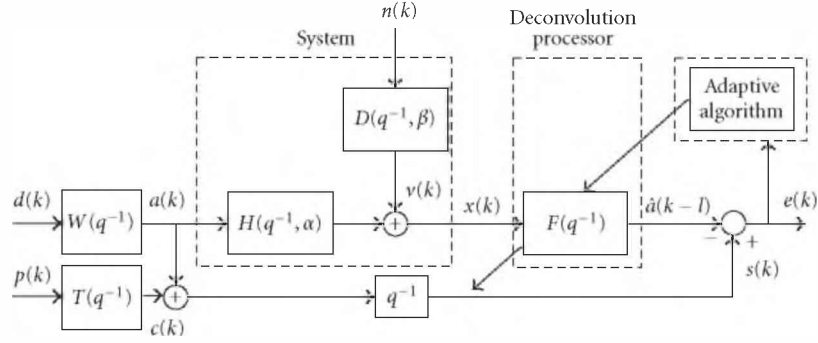
FIGURE 1: Block diagram of the general SISO deconvolution system including an adaptation algorithm.

it refers to the degree of the polynomial in negative powers of the associated variable. For proper or strictly proper rational functions, $\mathcal{O}(\cdot)$ is the degree of the denominator polynomial. For example, let H denote the rational function $H = H(q^{-1}, \boldsymbol{\alpha}) = (b_0 + b_1 q^{-1} + \cdots + b_M q^{-M})/(1 + a_1 q^{-1} + \cdots + a_N q^{-N})$, then the numerator of H is $\mathcal{O}(M, 0) = \mathcal{O}(M)$, the denominator is $\mathcal{O}(N, 0) = \mathcal{O}(N)$, and if $N \geq M$, H is $\mathcal{O}(N)$. In (1), H is of $\mathcal{O}(N)$, D is of $\mathcal{O}(S)$ and the shaping filters $W(q^{-1})$ and $T(q^{-1})$ are $\mathcal{O}(P)$ and $\mathcal{O}(V)$, respectively.

The signals $d(k)$, $p(k)$, and $n(k)$ play different roles depending on the particular application. In classical deconvolution, $p(k) = 0$ and $d(k)$ is colored by W to generate the input signal $a(k)$. The corrupting noise is represented by $D(q^{-1})n(k)$. In this case, F is designed as a linear processor that produces an efficient estimate of a possibly delayed version of the signal $a(k)$. Estimation is performed by linear filtering or smoothing operations on the noise-corrupted output signal of H, $x(k)$.

The signal enhancement problem can also be considered letting $p(k) \neq 0$. The signal of interest $p(k)$ (or $T(q^{-1})p(k)$), corrupted by the interference $a(k)$, is to be recovered by subtracting from $s(k)$ a filtered and noise-corrupted version of $a(k)$, that is, $x(k)$. The filters H and D are not completely known. The error $e(k)$ is actually the estimated value of $p(k)$. The goal is to design the linear processor F that will efficiently, in some well-defined sense, estimate the interference signal $a(k)$ (or $W(q^{-1})d(k)$).

Yet another application that is contemplated by the scheme of Figure 1 is the problem of linear equalization, which is described in detail in the example of Section 5.

All these deconvolution problems casted in the common framework of Figure 1 and described mathematically by (1) share the same formulation and solution, as will be shown later in this section.

2.2. System uncertainty description

The system uncertainties are modeled as

$$H(\boldsymbol{\alpha}) = H(\boldsymbol{\alpha}_0 + \boldsymbol{\delta}_\alpha) = H(\boldsymbol{\alpha}_0) + \Delta H,$$
$$D(\boldsymbol{\beta}) = D(\boldsymbol{\beta}_0 + \boldsymbol{\delta}_\beta) = D(\boldsymbol{\beta}_0) + \Delta D, \tag{3}$$

where $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{\delta}_\alpha$ and $\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \boldsymbol{\delta}_\beta$ are the parameters vectors with $\boldsymbol{\alpha}_0$ and $\boldsymbol{\beta}_0$ representing the nominal or mean value of the parameters. The vectors $\boldsymbol{\delta}_\alpha$ and $\boldsymbol{\delta}_\beta$ are independent zero

mean random perturbations, with a priori known covariance matrices $E[\boldsymbol{\delta}_\alpha \ \boldsymbol{\delta}_\alpha^T] = \boldsymbol{\gamma}_\alpha$ and $E[\boldsymbol{\delta}_\beta \ \boldsymbol{\delta}_\beta^T] = \boldsymbol{\gamma}_\beta$. The uncertainty on the parameters represented by $\boldsymbol{\delta}_\alpha$ and $\boldsymbol{\delta}_\beta$ results in an uncertain system which can be thought of as having different realizations for each particular value of the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, as shown by (3).

There are several approaches for the description of the additive perturbations ΔH and ΔD. These methods range from adjusting simple models to the set of systems from time or frequency experimental data, to the development of usually detailed high-order models that tightly describe the uncertainty boundaries in a certain range of frequencies of interest. See for example [17, 18, 19, 20]. The derivation in [20] could be of particular interest if a common orthogonal basis framework for the representation of the system, uncertainty and deconvolver, is pursued.

Without loss of generality, and keeping in mind the existence of more refined approaches, a simple linear approximation is adopted following a formulation close to that of Lin et al. in [6] or Chen and Lin in [7].

Expanding $H(\boldsymbol{\alpha})$ and $D(\boldsymbol{\beta})$ around the values $H(\boldsymbol{\alpha}_0)$ and $D(\boldsymbol{\beta}_0)$ in Taylor series and retaining the linear terms yields

$$\Delta H \approx (\boldsymbol{\delta\alpha})^T \frac{\partial H(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \bigg|_{\alpha = \alpha_0},$$
$$\Delta D \approx (\boldsymbol{\delta\beta})^T \frac{\partial D(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \bigg|_{\beta = \beta_0}, \tag{4}$$

where $\partial H(\alpha)/\partial\alpha$ and $\partial D(\beta)/\partial\beta$ are the Jacobian matrices of H and D, respectively. With the models (4), the statistical characterization of the system uncertainties is straightforward

$$E[\Delta H] = 0, \tag{5}$$

$$\Gamma_{\Delta H} = E\left[\Delta H^*(\boldsymbol{\alpha}) \ \Delta H(\boldsymbol{\alpha})\right]$$
$$= \left(\frac{\partial H(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\right)^*_{\alpha = \alpha_0} \boldsymbol{\gamma}_\alpha \left(\frac{\partial H(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\right)_{\alpha = \alpha_0}, \tag{6}$$

$$E[\Delta D] = 0, \tag{7}$$

$$\Gamma_{\Delta D} = E\left[\Delta D^*(\boldsymbol{\beta}) \ \Delta D(\boldsymbol{\beta})\right]$$
$$= \left(\frac{\partial D(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)^*_{\beta = \beta_0} \boldsymbol{\gamma}_\beta \left(\frac{\partial D(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)_{\beta = \beta_0}. \tag{8}$$

2.3. A family of orthogonal basis function

A generalized type of orthonormal construction will be used for the deconvolution processor F. Some of the advantages of this type of realizations for adaptive infinite impulse response (IIR) filters are discussed in [12].

If F is a linear time-invariant stable filter (or smoother), it can be expanded and represented as

$$F(q^{-1}) = \sum_{n=0}^{\infty} \theta_n L_n(q^{-1}, \Lambda_n) \tag{9}$$

with $L_n(q^{-1}, \Lambda_n)$ a complete set of orthonormal basis functions in the Hilbert space $\mathcal{H}_2$ of square (Lebesgue) integrable functions on the unit circle $\{z : |z| = 1\}$ and analytic for $|z| > 1$. These basis functions are characterized by the subset Λ_n of parameters taken from the general (finite or infinite) set

$$\Lambda = \{\lambda_0, \lambda_1, \ldots, \lambda_i, \ldots\} \tag{10}$$

with $\lambda_i \in \mathcal{C}$. The practical idea is to approximate F in (9) with a finite number of terms and a finite set of parameters from Λ. The following basis functions were reported in [14] and will be used in the expansion (9):

$$L_n(q^{-1}, \Lambda_F) = q^d \nu_n \frac{q^{-1}}{1 - q^{-1}\lambda_n} \prod_{k=0}^{n-1} \frac{q^{-1} - \lambda_k^*}{1 - q^{-1}\lambda_k}, \tag{11}$$

where $\nu_n = \sqrt{1 - |\lambda_n|^2}$ is the normalization constant, d is 0 or 1, and Λ_F is a finite set of parameters that depend on the function F. The functions in (11) have the property of allowing the inclusion of a variety of modes (different basis parameters usually coincident with the poles of F). Furthermore, they provide a unifying formulation for almost all known system identification orthonormal constructions such as FIR, Laguerre, and Kautz models. Moreover, methods using balanced realizations of user-chosen dynamics such as that presented in [21] can also be generated by (11). From a practical point of view, the inclusion of different modes means that F may be exactly represented by (9) and with a finite number of terms if the basis parameters are adequately chosen. Another relationship stemming from (11) and useful for implementation purposes is the recursive form

$$L_{n+1} = \eta_{n+1} L_n \frac{\overline{C}(\lambda_n)}{C(\lambda_{n+1})}, \tag{12}$$

where $\eta_{n+1} = \nu_{n+1}/\nu_n$, $C(\lambda_n) = 1 - q^{-1}\lambda_n$ and $\overline{C}(\lambda_n) = q^{-1} - \lambda_n$. Equations (11) and (12) are valid for real or complex parameters. Usually, in linear dynamical systems and for physical considerations, complex poles appear in conjugate pairs and the impulse response of the system is real. In this case, the new basis functions associated with the complex poles pairs are built in a different way. The construction uses

linear combinations of those generated by (11), preserving orthogonality and assuring a real-valued impulse response [14]. For each pair of complex poles then, and if $d = 0$, the associated basis has the form

$$
\begin{aligned}
&L_n'(q^{-1}, \Lambda) \\
&= \nu_n \frac{q^{-1}(a' + b'q^{-1})}{1 - (\lambda_n + \lambda_n^*)q^{-1} + |\lambda_n|^2} \prod_{k=0}^{n-1} \frac{q^{-1} - \lambda_k^*}{1 - q^{-1}\lambda_k}, \\
&L_n''(q^{-1}, \Lambda) \\
&= \nu_n \frac{q^{-1}(a'' + b''q^{-1})}{1 - (\lambda_n + \lambda_n^*)q^{-1} + |\lambda_n|^2} \prod_{k=0}^{n-1} \frac{q^{-1} - \lambda_k^*}{1 - q^{-1}\lambda_k},
\end{aligned}
\tag{13}
$$

where $\mathbf{x}_1 = [a' \ b']^T$ is chosen to belong to $\mathbf{x}_1^T \mathbf{M} \mathbf{x}_1 = |1 - \lambda_n^2|^2$ with

$$\mathbf{M} = \begin{bmatrix} 1 + |\lambda_n|^2 & 2\,\mathrm{Re}\{\lambda_n\} \\ 2\,\mathrm{Re}\{\lambda_n\} & 1 + |\lambda_n|^2 \end{bmatrix}. \tag{14}$$

The other pair of coefficients grouped by vector $\mathbf{x}_2 = [a'' \ b'']^T$ can then be found as a function of $\mathbf{x}_1$ by evaluating

$$\mathbf{x}_2 = \frac{1}{\sqrt{1 - \rho^2}} \begin{bmatrix} \rho & 1 \\ -1 & -\rho \end{bmatrix} \mathbf{x}_1, \tag{15}$$

where $\rho = (\lambda_n + \lambda_n^*)/(1 + |\lambda_n|^2)$. With these expressions, and if the components of Λ_F are real or complex conjugate pairs, the basis functions will have real impulse responses.

2.4. Problem formulation

From Figure 1, using (1), and with the system model given by (3), the error sequence $e(k)$ is

$$
\begin{aligned}
e(k) = &[d(k)W(q^{-l} - H(\boldsymbol{\alpha}_0)F) - n(k)D(\boldsymbol{\beta}_0)F + p(k)Tq^{-l}] \\
&- [d(k)W\Delta H + n(k)\Delta D]F.
\end{aligned}
\tag{16}
$$

Assuming the signals $d(k)$, $p(k)$, and $n(k)$ are also statistically independent of the model uncertainties and using (5)–(8), the *MSE over the models uncertainties* becomes

$$
\begin{aligned}
\xi &= E_\Delta[e^*(k)e(k)] \\
&= [d(k)W(q^{-l} - H(\boldsymbol{\alpha}_0)F) - n(k)D(\boldsymbol{\beta}_0)F + p(k)Tq^{-l}]^* \\
&\quad \times [d(k)W(q^{-l} - H(\boldsymbol{\alpha}_0)F) - n(k)D(\boldsymbol{\beta}_0)F + p(k)Tq^{-l}] \\
&\quad + W^*d^*(k)F^*\Gamma_{\Delta H}Fd(k)W + n^*(k)F^*\Gamma_{\Delta D}Fn(k),
\end{aligned}
\tag{17}
$$

where the operator $E_\Delta[\cdot]$ is the expectation applied only over the uncertainties in the models ΔH and ΔD. As a measure of

performance, the mean value of ξ over time $E_k[\cdot]$ is considered, and this is simply the *MSE*,

$$
\begin{aligned}
J(F) &= E_k[\xi] \\
&= \sigma_d^2 W^* W \left[1 - (q^{-l})^* HF - F^* H^* q^{-l} \right] \\
&\quad + F^* \psi^* \psi F + T^* T \sigma_{\tilde{p}}^2,
\end{aligned}
\tag{18}
$$

where ψ is the *minimum phase right spectral factor* of the spectral factorization [22]

$$
\psi^* \psi = \sigma_d^2 W^* W (H^* H + \Gamma_{\Delta H}) + \sigma_n^2 (D^* D + \Gamma_{\Delta D}).
\tag{19}
$$

Taking into account the general objective of designing a deconvolution processor robust to parameter uncertainty with an orthogonal structure, the problem formulation may now be summarized in the following statement. *Given the system (1), find the causal and stable deconvolution processor F_o, with the structure given by (9) and using the orthonormal functions (11), that minimizes the performance index J of (18).*

3. PROBLEM SOLUTION

Theorem 1. *For the system (1), the optimal causal and stable deconvolution processor with the orthogonal structure of (9) that minimizes the performance index J given by (18) is*

$$
F_o = \sum_{n=0}^{2(N+S+P)+l} \theta_{on} L_n(z^{-1}, \Lambda_o).
\tag{20}
$$

The maximum number of terms of (20) is $M = 2(N+S+P) + l + 1$ and Λ_o is the optimal basis parameter set,

$$
\Lambda_o = \{\lambda_z, \lambda_\psi\}
\tag{21}
$$

with $\lambda_z = \{0, \ldots, 0, p_{W_1}, \ldots, p_{W_P}\}$, composed of $l+1$ zeros and P additional parameters, p_{W_i} that are the poles of W, and $\lambda_\psi = \{z_1, \ldots, z_{2(N+S)+P}\}$ where the z_i are the $2(N+S) + P$ zeros of ψ.
The optimal coefficients of (20) are

$$
\Theta_o = [\theta_{o0}, \theta_{o1}, \ldots, \theta_{o2(N+S)+P+l}]^T
\tag{22}
$$

with

$$
\theta_{on} = \frac{1}{2\pi j} \oint L_n(z^{-1}, \Lambda_o) \left(\{Qz^{-l}\}_+ \psi^{-1} \right)^* z^{-1} dz,
\tag{23}
$$

$$
Q = \sigma_d^2 W^* W H^* (\psi^*)^{-1}.
\tag{24}
$$

Proof. See the appendix. □

3.1. Comments on these results

This theorem establishes the parameters Λ_o and the coefficients Θ_o that completely define the deconvolver F_o with the orthogonal structure given by (20), together with the maximum number of basis functions required. In this sense, the theorem solves one of the problems usually associated with the approximation of functions with orthogonal basis, which is the way the parameters have to be chosen to optimally approximate a desired function [23]. In this case the desired function is the optimal deconvolution processor and the representation achieved using the bases is exact, it is not an approximation. This is so because of the multiple modes (parameters or poles) admissible by the basis functions. Also, these sets of parameters and coefficients represent the best choice in the MSE sense that defines a deconvolver capable of dealing with a whole family of systems as described by (3) and (4). Again, in this sense, we say the orthogonal deconvolver is robust to parameter uncertainty in the system.

The poles of the orthogonal deconvolver are defined by Λ_o in (21). This set is composed by $l+1$ poles in zero plus the poles of W plus the zeros of ψ. It can be directly verified that in the case when no noise is present ($n(k) = 0$ or $D = 0$), the input is white ($W = 1$), the delay $l = 0$, H is minimum phase, and the parameters are unperturbed, then $F_o = H^{-1}$ and Λ_o just groups the zeros of H. For this case, the coefficients Θ_o will be such that the zeros of the numerator of the rational function resulting from (20) are the poles of H.

In the appendix during the proof of the theorem the following expression appears as an intermediate result for the optimal deconvolution processor: $F_o = \{Qz^{-l}\}_+ \psi^{-1}$. This expression is coincident with that obtained in [7] and may be compared with the classical Wiener filtering results, for example, in [24]. It is particularly useful to analyze and interpret some of the characteristics of the optimal deconvolver that finally appear in the orthogonal structure. First, F_o may be considered as a cascade of two filters. The filter ψ^{-1} has an inherent recursive structure that is independent of the delay (see in (19) that ψ is fixed and unique for a given system and shaping filter W). From (24), the filter $\{Qz^{-l}\}_+$ has the poles of W and $l+1$ poles in zero. When the design delay l changes, only this part of the deconvolution processor varies accordingly. When $W = 1$, that is, when the input is white noise, the deconvolution processor is a cascade of an FIR filter and an IIR filter. In this case, only the zeros part of λ_z will be present. So, if $W = 1$ and $l = 0$, the IIR part of the deconvolution processor is the optimal filter up to a scale factor. If $l > 0$, the deconvolver is a smoother and the FIR part of the processor performs the smoothing while the IIR portion remains unchanged. Any improvement in the performance of the deconvolution processor is generated by the FIR and the number of taps of this filter depends directly on the order of the delay l.

An additional comment applies referring to the structure of the deconvolver. The form of (20) is not the most practical from the point of view of implementation. Using the relation (12), the whole set of basis function can be generated as a cascade of first-order or second-order filters, depending on whether the poles are real or complex conjugate. This structure is illustrated in Figure 2 for the case when the basis parameters are real. It results in a very modular construction where additional basis functions can be easily incorporated if needed without affecting the existing structure.
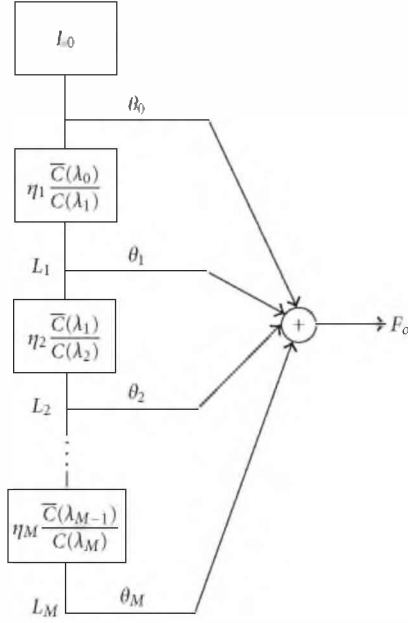
FIGURE 2: Practical structure for the optimal deconvolver, illustrated for $M + 1$ basis functions when the parameters λ_i are real.

3.2. Design algorithm

Before considering the incorporation of some adaptive capability to the deconvolver, the steps or algorithm for the optimal robust orthonormal design are summarized.

(1) Given the system and signal descriptions, choose the parameters that will be considered uncertain so as to give a good representation of the measured effects.
(2) Evaluate $\Gamma_{\Delta H}$ and $\Gamma_{\Delta D}$ with (6) and (8), respectively.
(3) Evaluate the spectral factorization (19).
(4) Evaluate (24).
(5) Evaluate the basis parameters (poles) of (21), that is, $l + 1$ zeros, plus the poles of W, plus the zeros of ψ, and build the basis.
(6) Evaluate the basis combining coefficients Θ with (23).
(7) The robust orthonormal deconvolution processor is built with (20) or using the equivalent representation based in the recursive expression (12) as shown for example in Figure 2.

The recursive form is preferred from the point of view of implementation and also convenient for the development of the adaptation strategy for the Θ.

4. COEFFICIENTS UPDATE

The robust orthogonal design can handle systems whose perturbation parameters δ_α and δ_β are small enough for the Taylor series expansion in (4) to remain valid. When the system departs from such region, the MSE performance deteriorates. In order to keep a desired performance for larger perturbations and also for tracking slowly time varying systems, some

degree of adaptivity is incorporated by updating only the coefficients of the linear combination of the basis functions. The main assumption is that the nominal or mean model for the system is still valid and representative of the real system and only the uncertainty region results enlarged. The basis structure remains fixed as well as the parameters Λ_o and the new set of coefficients Θ that now approximate the optimal deconvolver will be close to the initial optimal robust design. Figure 1 includes an updating algorithm in the general scheme of the deconvolver and Figure 3 illustrates the case when $W = 1$, $l > 0$, and Λ_o is real, so the deconvolver has the FIR-IIR cascade structure mentioned in the previous section with the coefficients Θ being updated by an adaptation algorithm.

4.1. Updating algorithm

The coefficients calculated from (23) are now treated as time varying and denoted accordingly as

$$\Theta = \Theta(k) = \left[\theta_0(k), \theta_1(k), \ldots, \theta_{2(N+S)+P+l}(k)\right]^T. \quad (25)$$

The updating algorithm is derived by minimizing an error functional $\varsigma(\Theta, k)$ that is a function of the coefficients,

$$
\begin{aligned}
\varsigma(\Theta, k) &= E\left\{[s(k) - \hat{a}(k - l)]^2\right\} \\
&= E\left\{[s(k) - \Theta^T(k)\mathbf{X}(k)]^2\right\},
\end{aligned}
\quad (26)
$$

where

$$\mathbf{X}(k) = \left[L_0, L_1, \ldots, L_l, L_{l+1}, \ldots, L_{2(N+S)+P+l}\right]^T x(k) \quad (27)$$

is a generalized regressor composed of the input signal to the deconvolution processor $x(k)$, filtered by the basis functions. Depending on the number of zeros in λ_z, the generalized regressor may include some delayed samples of $x(k)$, for example, in the case illustrated in Figure 3. Expanding (26),

$$\varsigma(\Theta, k) = E\{s^2(k)\} - 2\Theta^T(k)\mathbf{U}(k) + \Theta^T(k)\mathbf{R}_I(k)\Theta(k) \quad (28)$$

with $\mathbf{U}(k) = E\{s(k)\mathbf{X}(k)\}$ and $\mathbf{R}_I(k) = E\{\mathbf{X}(k)\mathbf{X}^T(k)\}$. A gradient-based family of adaptive algorithms can be generated by using a coefficient-updating equation of the form

$$\Theta(k + 1) = \Theta(k) - \mu\mathbf{G}(k), \quad (29)$$

where

$$\mathbf{G}(k) = \frac{\partial \varsigma(\Theta, k)}{\partial \Theta} = 2[\mathbf{R}_I(k)\Theta(k) - \mathbf{U}(k)] \quad (30)$$

is the gradient vector of the error functional (28) in the coefficients space and μ is the convergence factor, a small positive real number. Different approaches for the evaluation of an estimate of the real theoretical gradient $\mathbf{G}(k)$ result in different algorithms. One of the most popular approaches uses the instantaneous values of $\mathbf{U}(k)$ and $\mathbf{R}_I(k)$ as estimates of their
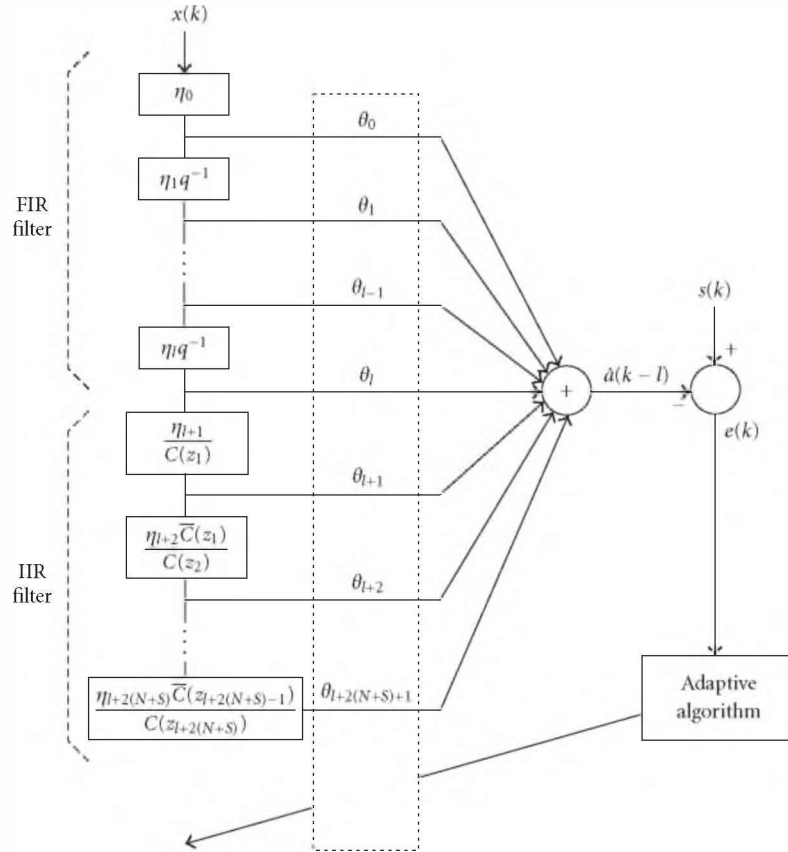
FIGURE 3: Structure of the orthogonal robust adaptive deconvolution processor for real basis parameters when $W = 1$ and $l > 0$.

means, that is,

$$\hat{\mathbf{U}}(k) = s(k)\mathbf{X}(k),$$
$$\hat{\mathbf{R}}_I(k) = \mathbf{X}(k)\mathbf{X}^T(k). \qquad (31)$$

Using (31) in the gradient (30),

$$\hat{\mathbf{G}}(k) = -2\mathbf{X}(k)e_I(k), \qquad (32)$$

where $e_I(k) = s(k) - \hat{a}(k - l)$ is the instantaneous error of the adaptive structure. Using this estimation for the gradient in (29), the equation for updating the coefficients is

$$\Theta(k + 1) = \Theta(k) + 2\mu e_I(k)\mathbf{X}(k) \qquad (33)$$

and the algorithm may be classified as a transform domain *least mean square* or LMS [25, 26]. With a slight increase in complexity, a recursive least squares or a lattice-like algorithm [27] may also be derived, but this will not be pursued here. The tracking capability and noise performance of this and other types of algorithms, related to these basis functions, have been analyzed in [13] for the application of system modeling. Also, issues related to convergence speed and other properties for orthogonal realizations of IIR filters were discussed in [12].

5. EXAMPLE: LINEAR ROBUST ADAPTIVE EQUALIZATION FOR AN ADSL TYPE OF COMMUNICATION CHANNEL

The general problem of equalization and particularly adaptive equalization is well described in [28] and a review with comparisons between recursive and nonrecursive techniques is given in [29]. Linear equalization is a particular case of the general deconvolution problem where $T = 0$. Additionally, the reference signal $s(k)$ (a delayed version of $a(k)$) is generated as the output of a decision device in the receiver, assuming the decisions are correct. Figure 4 illustrates the adaptive linear equalization setup. The parts of the diagram in dashed lines represent the practical implementation for the generation of the reference signal in the receiver. The following simplifying assumptions are made to design the equalizer for this example: the design delay is $l = 1$ and the data sequence is a white noise signal, $W = 1$. The modeling assumptions are discussed first, then the robust orthogonal design is shown and finally adaptation is considered. Performance comparisons are presented in these steps.

5.1. Modeling

Figure 5 shows the frequency response (FR, normalized to 0 dB at zero frequency) of a subscriber telephone loop, with a length of 2.9 Km (gauge 24 AWG) with a bridge tap of 100 meters of gauge 26 AWG, used in this example for asymmetric digital subscriber line (ADSL) transmissions. It
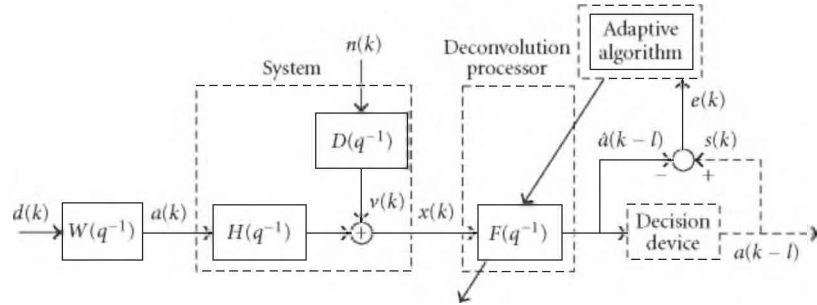
FIGURE 4: The setup of the general deconvolution system for the equalization problem. The outputs of the decision device are assumed to be correct.
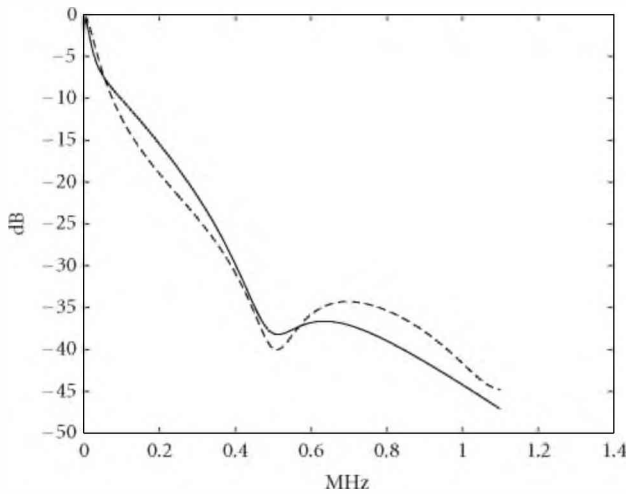

FIGURE 5: Real (solid line) and approximated (dashed line) frequency response of an ADSL loop.

was generated from the chain matrix characterization for this type of channels [30] with a bandwidth that extends to 1.104 Mhz.

The FR exhibits a notch around a frequency of 500 kHz. The frequency location of this notch is related to the minimum of the input impedance that presents an open circuited section of cable at frequencies for which the length is an odd number of quarter wavelengths. The attenuation or depth of the notch is proportional to the length and to the square root of the notch frequency. Also included in the same figure is the FR of a discrete third-order model designed to approximate the analog response. This model has the following expression in the transform domain:

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}}, \quad (34)$$

and is characterized by the nominal parameter vector

$\boldsymbol{\alpha}_0$

$= \begin{bmatrix} b_0 & b_1 & b_2 & b_3 & a_1 & a_2 & a_3 \end{bmatrix}^T$

$= \begin{bmatrix} 0.03 & 0.0153 & 0.0173 & 0.0171 & -1.0284 & 0.3307 & -0.2216 \end{bmatrix}^T.$

(35)

The response of this model is 4 dB within the real FR curve and it will be used for the purpose of illustrating the potential performance of the proposed linear deconvolver. Nevertheless, it should not be considered as a reference model for general ADSL systems or digital subscriber loops [31].

The effect of the variations of the individual numerator coefficients of H on the FR are illustrated in Figure 6. Perturbations on b_0 have important effects in the depth of the notch and the gain of the high frequency portion of the response. Changes in b_1 seem to affect the whole response in a rather mild way, preserving the basic shape and modifying the location of the notch. The coefficients b_2 and b_3 affect both the location and depth of the notch but do not have much influence in the low frequency portion of the response.

Although H is not a physical model and its parameters are not necessarily related to the loop parameters, the family of responses or channels generated by the changes in these parameters can be associated with the uncertainties that arise when attempting to describe the loop. Usually the length, the exact location of bridge taps, and the precise conformation of the loop are not known. Additionally, most parameters are indirectly determined by impedance measurements. All these facts add up and make the determination of the exact response of the channel a difficult task. Uncertainties arise naturally about the overall gain of the loop and the location and depth of the notch, even though the shape (or mean value) of the response will not suffer considerable changes. Thus, it seems reasonable to consider an uncertain description for the channel as follows. The model (34) represents the nominal channel and b_1 the perturbed parameter. In this way, variations in b_1 model potential uncertainties, without distorting the basic shape of the FR over the whole range of frequencies of interest.

One of the most severe types of interference in ADSL is the near-end crosstalk (NEXT) produced by the voltages and currents induced in the line by nearby pairs of wires [30, 32]. The "average and asymptotic" NEXT power is proportional to $f^{1.5}$ and depends on some parameters of the particular line. A first-order ARMA model $D = (d_0 + d_1 z^{-1})/(1 + c_1 z^{-1})$ is used to shape the white noise sequence $n(k)$ with a power spectrum similar to the NEXT interference. This filter is
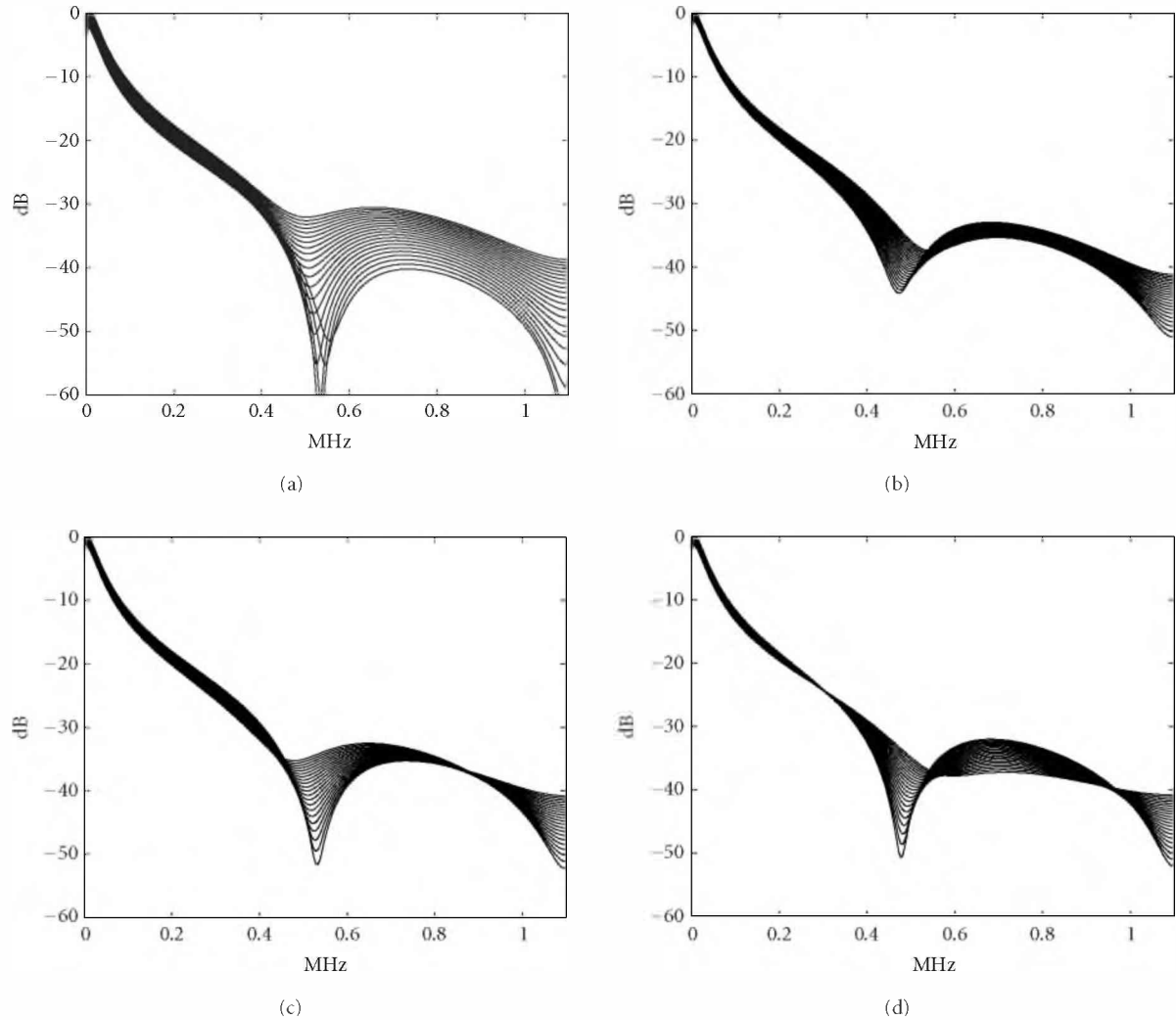
(a)

(b)

(c)

(d)

FIGURE 6: Frequency response of H when the numerator coefficients are perturbed. (a) Coefficient b_0. (b) Coefficient b_1. (c) Coefficient b_2. (d) Coefficient b_3.

characterized by the parameter vector

$$\boldsymbol{\beta}_0 = [d_0 \quad d_1 \quad c_1]^T = [0.0020 \quad -0.00196 \quad 0.7209]^T. \quad (36)$$

To control the signal-to-noise ratio (SNR) at the input of the equalizer, the variance or power of the signal measured at the output of the channel H, σ_y^2 is normalized to 1, and the gain of filter D is set in accordance with the following definition:

$$\text{SNR} = 10\log\left(\frac{\sigma_y^2}{\sigma_v^2}\right) = 10\log\left(\frac{1}{\sigma_v^2}\right), \quad (37)$$

where σ_v^2 is the variance of the colored noise at the output of D.

For adaptive equalization, transversal FIR filters are the standard choice for many reasons [22, 27, 28, 33], so comparisons with classical fixed recursive and adaptive FIR designs

are made. First, the number of coefficients required for an FIR equalizer will be evaluated. **Figure 7** shows the minimum MSE (MMSE) attainable as a function of the number of taps used for the equalizer. The family of curves is parameterized by the SNR. The MSE is limited by the SNR, so for low SNR, the performance of the equalizer is necessarily poor and only a few coefficients in the FIR are enough to attain the optimal performance. As the SNR rises, the number of taps needed to reach the MMSE is larger. If an SNR of 80 dB is considered the "no-noise design," then a minimum of 50 taps will be required by the FIR to approximate the optimal response.

5.2. Robust orthogonal design

Under the same design conditions, a similar analysis can be performed for the robust equalizer using the variance of the uncertain parameter b_1 as a "tuning knob." Figure 8 shows the MMSE attainable with the robust equalizer as a function of the SNR. The curves are parameterized by the variance $\sigma_{b_1}^2$.
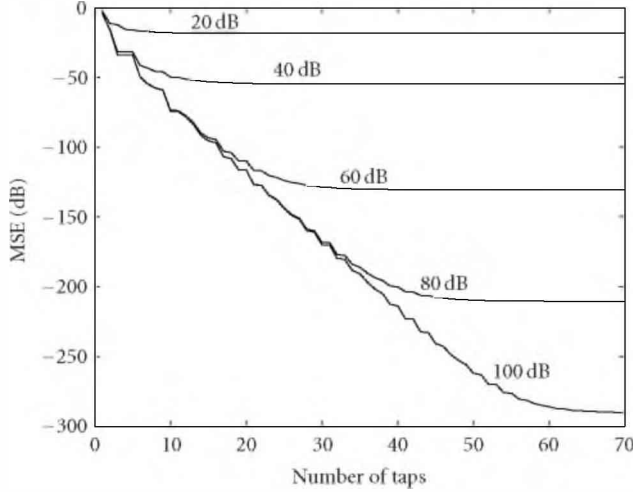
FIGURE 7: FIR equalizer. MSE as a function of the number of taps of the FIR. The parameter of the curves is the SNR.
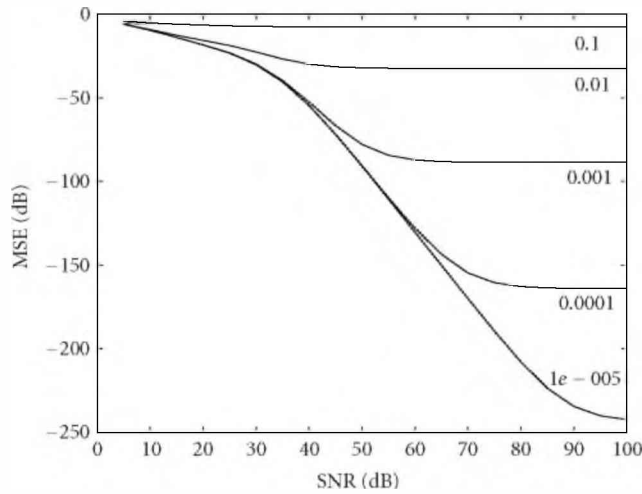

FIGURE 8: Robust equalizer. MSE as a function of the SNR at the input of the equalizer. The parameter of the curves is the variance of the coefficient b_1.

For low SNR, even the unperturbed IIR design (the lower curve for $\sigma_{b_1}^2 = 0.00001$ is almost coincident with the unperturbed design) has a poor performance with an MSE that is nearly in a one-to-one relation with the SNR. The curves show that the design variance has to be below 0.001 to obtain an MSE that is under -100 dB, that is, to obtain a performance similar to the FIR for the "no-noise design."

The effect of the variance of the parameter b_1 in the design may be better appreciated in Figure 9 that illustrates the MSE when the parameter b_1 departs from its nominal value for an SNR of 35 dB. The solid line curves correspond to the fixed nominal (unperturbed) IIR and 50-tap FIR designs. This two curves overlap, confirming that the FIR filter can very well approximate the optimal recursive equal-
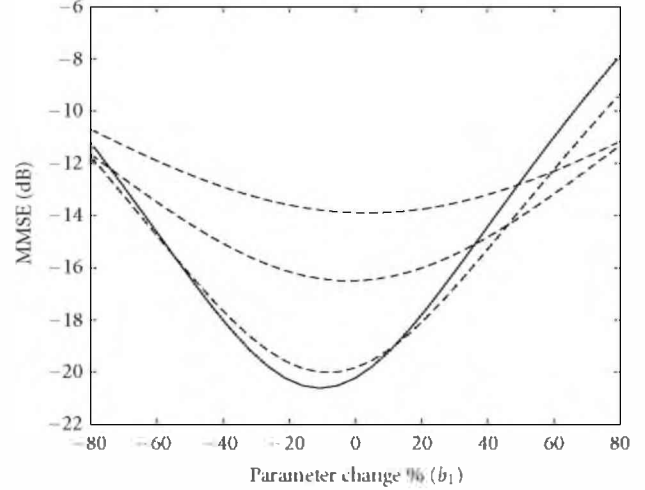

FIGURE 9: MSE versus percentage of variation of channel parameter b_1. Solid line: fixed unperturbed IIR and 50 taps FIR designs (the curves overlap). Dashed lines: robust designs for values of $\sigma_{b_1}^2$ of 0.001 (lower curve), 0.005, and 0.009 (upper curves).

izer. The dashed-line curves correspond to robust designs for different values of $\sigma_{b_1}^2$ (the lower error curve corresponds to $\sigma_{b_1}^2 = 0.001$). For higher variances, the designs are more conservative, the MSE grows and the curves tend to be "flatter." The performance is worst around the nominal value of the parameter but improves and even exceeds the nominal designs for larger deviations of b_1. This is very reasonable since robustness against channel uncertainty is obtained at the expense of lack of performance at the nominal value. These curves can be directly compared and coincide with those obtained using the approach of [7].

From the previous analysis we select $\sigma_{b_1}^2 = 0.001$, and the steps of the design algorithm for a SNR of 35 dB are as follows.

(1) The gains of H and D are adjusted according to (37) for an SNR of 35 dB considering $\sigma_a^2 = \sigma_d^2 = 1$ and $\sigma_n^2 = 1$,

$$H(z) = \frac{nH(z)}{dH(z)}$$

$$= \frac{0.1644 + 0.0839z^{-1} + 0.0947z^{-2} + 0.0936z^{-3}}{1 - 1.0284z^{-1} + 0.3307z^{-2} - 0.2216z^{-3}}, \quad (38)$$

$$D(z) = \frac{nD(z)}{dD(z)} = \frac{0.0067 - 0.0066z^{-1}}{1 + 0.7209z^{-1}}.$$

(2)

$$\Gamma_{\Delta H} = \frac{0.001}{(dH(z))^*(dH(z))}, \qquad \Gamma_{\Delta D} = 0. \quad (39)$$

(3)

$$\psi = \frac{0.1790 + 0.2003z^{-1} + 0.1567z^{-2} + 0.1552z^{-3} + 0.0620z^{-4}}{1.0000 - 0.3075z^{-1} - 0.4106z^{-2} + 0.0169z^{-3} - 0.1597z^{-4}}. \quad (40)$$

(4)

$$Q = \frac{0.1644 + 0.0334z + 0.0013z^2 + 0.0328z^3 - 0.0925z^4 - 0.0502z^5 - 0.0135z^6 - 0.0149z^7}{0.1790 + 0.0162z + 0.0098z^2 + 0.0207z^3 - 0.0902z^4 - 0.0471z^5 - 0.0139z^6 - 0.0137z^7}. \tag{41}$$

(5) The optimal basis parameters are

$$\Lambda_o = \{0, 0, -0.6797 \pm 0.1312i, 0.1202 \pm 0.8416i\}. \tag{42}$$

(6) The coefficients Θ_o are

$$\Theta_o = \begin{bmatrix} 0.5768 \\ 4.3072 \\ -6.3890 \\ 2.6193 \\ -3.6522 \\ 1.3904 \end{bmatrix}. \tag{43}$$

(7) The robust orthogonal equalizer is completely specified by (42) and (43) when using the functions (11) in (20) or in (12).

5.3. Robust and adaptive design

The previous design procedure did not incorporate coefficient adaptation. When adaptive equalization is considered, the FIR and the robust orthogonal design coefficients Θ are updated by an adaptation algorithm. The performance of the equalizers is evaluated in terms of the MMSE attainable when b_1 changes slowly with time (when compared to the convergence speed of the algorithms) and around the nominal design value. Figures 10 and 11 depict these results for two different SNRs. In Figure 10 the SNR is of 35 dB and the MMSE for the nominal design (solid line) is around -20 dB. The adaptive FIR is plotted with a dashed line and exhibits a great improvement in the performance when compared to fixed designs. This improvement is obtained at the cost of adapting all 50 coefficients. The dashed-dotted line corresponds to the robust adaptive orthogonal design that also improves the performance of the robust designs of Figure 9. It has almost the same performance as the FIR for positive perturbations of b_1, but is over 4 dB above for negative variations of more than 50% in the channel parameter. The main structure of the equalizer is fixed and only 6 coefficients are updated to obtain this performance.

Figure 10 also includes the response of a deconvolver designed as detailed in the previous sections but with an over-parameterized orthogonal basis representation (dotted line). Additional parameters are added to the optimal Λ_o to improve the performance for large deviations in b_1. The coefficients θ_i associated with these new parameters are almost zero when the perturbations are small and start to have significant values for larger departures. The optimal selection of

these additional parameters is related in this particular case to the zeros of H, and more generally to the zeros of ψ that change as the system is perturbed. In this example, the added parameters are $\{0.2225 \pm 0.9045i\}$. This means that the total number of coefficients to be adapted is 8 and the performance is almost the same as for the 50-tap adaptive FIR in the whole range of variation of b_1.

Figure 11 shows the MSE when the SNR is 50 dB. Again the performance of the robust adaptive orthogonal designs approaches the totally adaptive FIR, with only 6 to 8 adaptive coefficients.

5.4. Remarks

The example was developed assuming that the uncertainty is described by only one perturbed parameter just for clarity and simplicity of tuning. The procedure may be applied similarly, when more than one parameter is perturbed, including the coefficients of the denominator of H and D.

Over-parameterization of the basis functions may give significant improvement in the performance with little additional cost. A technique to find an optimal and systematic procedure for the selection of the additional parameters is actually a subject of research along with the potential problems of this type of over-parameterized adaptive and recursive structures [12, 25].

Possible extensions of the orthogonal adaptive structure to more specific applications in communications include the design of decision feedback equalizers. The feedforward and feedback filters of the DFE can be given an orthogonal structure with the basis (11) and a coefficient-updating strategy, similar to that of Section 4, used to make both filters partially adaptive. The feasibility of this approach was initially investigated in [34] and could be used for comparisons with the robust fixed designs of Lin et al. in [6] or Sternad et al. in [4], since both of these approaches deal with the problem of robust DFE design. Also in this area, the partially adaptive recursive structure for the feedback filter of the DFE may be a good alternative to long FIR adaptive filters [11, 31, 35]. Other applications and performance analysis of this approach are currently the subject of further research.

6. CONCLUSIONS

A design strategy for a general SISO robust orthogonal adaptive deconvolution processor has been presented. The approach reformulates and combines results from the fields of robust filtering, function approximation with orthogonal basis, and adaptive filtering. The design exhibits several of the
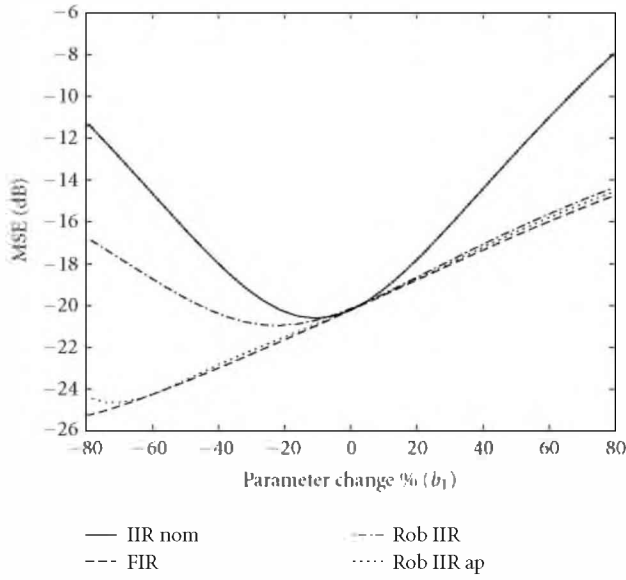
FIGURE 10: MSE versus variations in channel parameter b_1 for an SNR of 35 dB. Solid line: fixed nominal unperturbed IIR. Dashed line: adaptive FIR. Dashed-dotted line: robust orthogonal adaptive IIR. Dotted line: robust orthogonal adaptive IIR with additional basis parameters.

FIGURE 11: MSE versus variations in channel parameter b_1 for an SNR of 50 dB. Solid line: fixed nominal unperturbed IIR. Dashed line: adaptive FIR. Dashed-dotted line: robust orthogonal adaptive IIR. Dotted line: robust orthogonal adaptive IIR with additional basis parameters.

advantages related to these fields: (i) it is robust to parameter uncertainties in the system model; (ii) it is recursive and will require a smaller number of parameters than FIR counterparts for similar performance, hence the total computational burden is also smaller than for adaptive FIR designs; (iii) it has an orthogonal structure with good numerical properties and is very modular from an implementation point of view; (iv) it is adaptive and recursive but with a fixed pole structure so it does not have the potential stability problems of adaptive IIR filters [25, 27]; and (v) the complexity of the algorithms used for updating the coefficients is comparable to those used for FIR adaptive filters.

The main result was presented in the form of a theorem that puts together the design of the robust recursive deconvolver with an orthogonal basis representation and establishes how the basis parameters have to be selected.

An example was presented with the design of an equalizer for a simple ADSL channel model with NEXT interference. The simulation results show that the proposed design can extend the range of operation of fixed linear designs. It performs as well as FIR designs which require much more adaptive coefficients to yield acceptable results. Moreover, it was shown that the performance can be further improved by an over-parameterization of the orthogonal basis with a small increase in the number of adaptive parameters.

We summarize our contribution as having presented a design procedure for a filtering structure with a good trade-off between computational burden and performance, under a wide variety of conditions and uncertainties with applications throughout the signal processing, communications, and control fields.
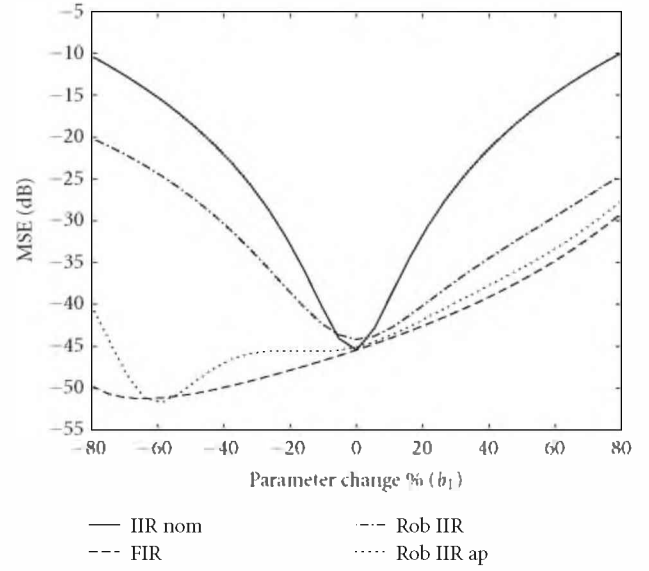
APPENDIX

PROOF OF THEOREM 1

The proof proceeds in two steps. First, a general expression for the optimal deconvolver is obtained, and second, the exact representation of this processor by means of the orthogonal basis is developed. For the first step, the minimization will be performed using the calculus of variations methodology [36] following a procedure close to that of [6, 7].

Initially a perturbation to the optimal unknown processor F_o is included in the following form:

$$F = F_o + \kappa \zeta(q^{-1}), \qquad (A.1)$$

where $\zeta(q^{-1})$ is an arbitrary, rational in q^{-1}, and realizable function, analytic on and outside the unit circle, and κ is a small bounded real constant. Replacing (A.1) in (18) and using Parseval's theorem to express the performance index in the transform domain,

$$J = \frac{1}{2\pi j} \oint_{|z|=1} [\sigma_d^2 W^* W (1 - z^l H F_o - F_o^* H^* z^{-l})$$
$$+ F_o^* \psi^* \psi F_o + \sigma_p^2 T^* T] \frac{dz}{z}$$
$$+ \frac{\kappa}{2\pi j} \oint_{|z|=1} [F_o^* \psi^* \psi \zeta + \zeta^* \psi^* \psi F_o$$
$$- \sigma_d^2 W^* W z^l H \zeta - \sigma_d^2 W^* W \zeta^* H^* z^{-l}] \frac{dz}{z}$$
$$+ \frac{\kappa^2}{2\pi j} \oint_{|z|=1} [\zeta^* \psi^* \psi \zeta] \frac{dz}{z}.$$

$$(A.2)$$

The simultaneous necessary and sufficient conditions to be satisfied for a minimum in J are [36]

$$\frac{\partial J}{\partial \kappa}\bigg|_{\kappa=0} = 0, \tag{A.3}$$

$$\frac{\partial^2 J}{\partial \kappa^2}\bigg|_{\kappa=0} > 0. \tag{A.4}$$

The conditions imposed by (A.4) mean that the integral

$$\frac{1}{2\pi j} \oint_{|z|=1} [\zeta^* \psi^* \psi \zeta] \frac{dz}{z} \tag{A.5}$$

must be positive, and this is always satisfied because the integrands involved have the symmetric form $f^*(z^{-1})f(z^{-1})$ and are always greater than zero when integrated over the unit circle. The condition of (A.3) implies that the following integral must be equal to zero:

$$\frac{1}{2\pi j} \oint_{|z|=1} \left[(F_o \psi \psi^* \zeta^* - \sigma_d^2 W^* W \zeta^* H^* z^{-l}) \right. \\ \left. + (F_o^* \psi^* \psi \zeta - \sigma_d^2 W^* W \zeta H z^l) \right] \frac{dz}{z} = 0. \tag{A.6}$$

Defining $Q = \sigma_d^2 W^* W H^* (\psi^*)^{-1}$, applying the $\{\cdot\}_+$ and $\{\cdot\}_-$ operators, and writing condition (A.6) as two integrals give the following equivalent condition:

$$\frac{1}{2\pi j} \oint_{|z|=1} \left\{ \psi^* \zeta^* \left[F_o \psi - \{Qz^{-l}\}_+ - \{Qz^{-l}\}_- \right] \right\} \frac{dz}{z} \\ + \frac{1}{2\pi j} \oint_{|z|=1} \left\{ \psi^* \zeta^* \left[F_o \psi - \{Qz^{-l}\}_+ - \{Qz^{-l}\}_- \right] \right\}^* \frac{dz}{z} = 0. \tag{A.7}$$

Applying Cauchy's theorem, (A.7) is satisfied if the following part of the integrand is zero,

$$F_o \psi - \{Qz^{-l}\}_+ = 0, \tag{A.8}$$

since it is the only one which may have poles inside the unit circle. Finally, from (A.8) the expression for the optimal robust deconvolution processor is

$$F_o = \{Qz^{-l}\}_+ \psi^{-1}. \tag{A.9}$$

Three comments on this result follow. First, a realizable F_o can only eliminate the parts of the integrand that involve $\{\cdot\}_+$ terms (the filter is analytic outside the unit circle and as so, causal). Second, the $\{\cdot\}_-$ term is a rational function starting with a free z to cancel the pole at the origin of the integrand in (A.7). Third, for symmetry reasons, if one of the integrals in (A.7) is zero, so will be the other. This derivation is usually followed in the classical Wiener filtering approach (see for example [37]), and the reader may compare this result with the ones in [6, 7, 24].

We now perform an analysis of the operations involved in (A.9) to establish bounds on the maximum degree of the polynomials that conform the optimal deconvolver and justify the parameter assignment of the orthogonal basis.

Recalling that H and D are up to $\mathcal{O}(N)$ and $\mathcal{O}(S)$, respectively, and assuming that at least one of the denominator parameters of these functions is perturbed, then, from (6), $\Gamma_{\Delta H}$ is rational with numerator polynomial of $\mathcal{O}(N,N)$ and denominator $\mathcal{O}(2N,2N)$. For $\Gamma_{\Delta D}$, the degrees are $\mathcal{O}(S,S)$ for the numerator and $\mathcal{O}(2S,2S)$ for the denominator, respectively. Both polynomials of the rational spectral factorization ψ^* of (19) have degrees upper bounded by $\mathcal{O}(2N+2S+P,2N+2S+P)$. Performing the product to calculate Qz^{-l} results in a rational function of $\mathcal{O}(P+l,3N+2S+P)$ for both numerator and denominator polynomials. All the poles of Q are outside the unit circle, except those afforded by W of $\mathcal{O}(P)$. Thus, $\{Qz^{-l}\}_+$ is $\mathcal{O}(P)$ if $l=0$, or $\mathcal{O}(P+l)$ for the numerator polynomial if $l>0$. The optimal deconvolver is then conformed by a cascade of two filters. One is ψ^{-1} of maximum $\mathcal{O}(2N+2S+P)$. The other is $\{Qz^{-l}\}_+$ with a maximum number of terms for the numerator polynomial of $P+l$.

To represent exactly the optimal deconvolver by means of the basis functions (11), the parameter set Λ has to be assigned to match the pole structure of F_o. This means setting the P poles of W, plus $l+1$ poles at the origin to account for the delay, plus $2N+2S+P$ additional poles that are the zeros of ψ. This justifies the parameter assignment of (21) and the number of terms involved in (20).

Finally, (23) is the standard inner product used to determine the coefficients that linearly combine the basis functions once the parameters have been optimally assigned. This concludes the proof.

ACKNOWLEDGMENT

REFERENCES

[1] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, 1998.

[2] J. F. Cardoso, "Blind signal separation: statistical principles," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009–2025, 1998.

[3] M. Sternad and A. Ahlén, "Robust filtering and feedforward control based on probabilistic descriptions of model errors," *Automatica*, vol. 29, no. 3, pp. 661–679, 1993.

[4] M. Sternad, A. Ahlén, and E. Lindskog, "Robust decision feedback equalizers," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '93)*, vol. 3, pp. 555–558, Minneapolis, Minn, USA, April 1993.

[5] M. Sternad and A. Ahlén, "$\mathcal{H}_2$ design of nominal and robust discrete time filters," in *A Polynomial Approach to H_2 and H_∞ Robust Control Design*, M. Grimble and V. Kučera, Eds., pp. 171–222, Springer-Verlag, London, UK, 1996.

[6] C.-W. Lin, B.-S. Chen, and T.-Y. Yang, "Robust decision-feedback equalizer design method under parameter perturbation in a transmission channel," *Signal Processing*, vol. 44, no. 1, pp. 39–49, 1995.

[7] B.-S. Chen and C.-W. Lin, "Optimal deconvolution filter design under parameters perturbation in transmission channels," *Circuits, Systems, and Signal Processing*, vol. 15, no. 2, pp. 213–231, 1996.

[8] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "The fundamental role of general orthonormal bases in system identification," *IEEE Trans. Automatic Control*, vol. 44, no. 7, pp. 1384–1406, 1999.

[9] B. Ninness, S. Gibson, and S. Weller, "Practical aspects of using orthonormal system parameterisations in estimation problems," in *Proc. 12th IFAC Symposium on System Identification*, vol. 2, pp. 463–468, Santa Barbara, Calif, USA, June 2000.

[10] Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Trans. Speech and Audio Processing*, vol. 2, no. 2, pp. 320–328, 1994.

[11] N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: a unified approach," *IEEE Transactions on Information Theory*, vol. 42, no. 3, pp. 903–915, 1996.

[12] J. Cousseau, P. S. R. Diniz, G. Sentoni, and O. Agamennoni, "On orthogonal realizations for adaptive IIR filters," *International Journal Circuit Theory and Applications*, vol. 28, no. 5, pp. 481–500, 2000.

[13] B. Ninness and J. C. Gómez, "Frequency domain analysis of tracking and noise performance of adaptive algorithms," *IEEE Trans. Signal Processing*, vol. 46, no. 5, pp. 1314–1332, 1998.

[14] B. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Trans. Automatic Control*, vol. 42, no. 4, pp. 515–521, 1997.

[15] P. Doñate, C. Muravchik, and J. Cousseau, "Robust orthogonal adaptive deconvolution," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS '01)*, vol. 2, pp. 733–736, Sydney, NSW, Australia, May 2001.

[16] G. W. Davidson and D. D. Falconer, "Reduced complexity echo cancellation using orthonormal functions," *IEEE Trans. Circuits and Systems*, vol. 38, no. 1, pp. 20–28, 1991.

[17] H. Hindi, C.-Y. Seong, and S. Boyd, "Computing optimal uncertainty models from frequency domain data," in *proc. 41st IEEE Conference on Decision and Control*, vol. 3, pp. 2898–2905, Las Vegas, Nev, USA, December 2002.

[18] A. Vicino and G. Zappa, "Sequential approximation of parameter sets for identification with parametric and nonparametric uncertainty," in *proc. 32nd IEEE Conference on Decision and Control*, vol. 3, pp. 2044–2049, San Antonio, Tex, USA, December 1993.

[19] M. E. Campbell and S. C. O. Grocott, "Parametric uncertainty model for control design and analysis," *IEEE Transactions on Control Systems Technology*, vol. 7, no. 1, pp. 85–96, 1999.

[20] P. Doñate, C. Muravchik, and J. Cousseau, "Modeling and decision feedback equalization of uncertain channels," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS '99)*, vol. 4, pp. 536–539, Orlando, Fla, USA, May–June 1999.

[21] P. S. C. Heuberger, P. M. J. Van den Hof, and O. H. Bosgra, "A generalized orthonormal basis for linear dynamical systems," *IEEE Trans. Automatic Control*, vol. 40, no. 3, pp. 451–465, 1995.

[22] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic, Boston, Mass, USA, 2nd edition, 1994.

[23] P. Bodin, L. F. Villemoes, and B. Wahlberg, "An algorithm for selection of best orthonormal rational basis," in *Proc. 36th IEEE Conference on Decision and Control*, vol. 2, pp. 1277–1282, San Diego, Calif, USA, December 1997.

[24] A. Ahlén and M. Sternad, "Wiener filter design using polynomial equations," *IEEE Trans. Signal Processing*, vol. 39, no. 11, pp. 2387–2399, 1991.

[25] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementations*, Kluwer Academic, Boston, Mass, USA, 1997.

[26] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1985.

[27] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 3rd edition, 1996.

[28] S. U. H. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, vol. 73, no. 9, pp. 1349–1387, 1985.

[29] N. J. Lynch-Aird, "Review and analytical comparison of recursive and nonrecursive equalization techniques for PAM transmission systems," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 6, pp. 830–838, 1991.

[30] J. A. C. Bingham, *ADSL, VDSL, and Multicarrier Modulation*, chapter 3, J. G. Proakis Ed., John Wiley & Sons, New York, NY, USA, 2000.

[31] N. Al-Dhahir, A. H. Sayed, and J. M. Cioffi, "Stable pole-zero modeling of long FIR filters with application to the MMSE-DFE," *IEEE Trans. Communications*, vol. 45, no. 5, pp. 508–513, 1997.

[32] S. H. Lin, "Statistical behavior of multipair crosstalk," *Bell System Technical Journal*, vol. 59, no. 6, pp. 955–974, 1980.

[33] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 2nd edition, 1989.

[34] P. Doñate, C. Muravchik, and J. Cousseau, "On the design of robust orthogonal adaptive decision feedback equalizers for uncertain dispersive channels," *Latin American Applied Research*, vol. 33, no. 4, pp. 419–425, 2003.

[35] N. Al-Dhahir, "Time-varying versus time-invariant finite-length MMSE-DFE on stationary dispersive channels," *IEEE Trans. Communications*, vol. 46, no. 1, pp. 11–15, 1998.

[36] G. Leitmann, *The Calculus of Variations and Optimal Control*, Plenum Press, New York, NY, USA, 1981.

[37] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, John Wiley & Sons, New York, NY, USA, 1968.

P. D. Doñate was born in Bahía Blanca, Argentina, 1958. He graduated as an electrical engineer from the Universidad Nacional del Sur, Bahía Blanca, Argentina, in 1983. He is Professor of circuits and systems, communication systems, and estimation theory at the Department of Electrical and Computer Engineering, working in the Signal Processing and Communications Laboratory. His research interests are in the areas of signal processing and recursive filtering with applications in control and communication systems.

C. Muravchik was born in Argentina, June 11, 1951. He graduated as an electronics engineer from the National University of La Plata, Argentina, in 1973. He received the M.S. in statistics (1983) and the M.S. (1980) and Ph.D. (1983) degrees in electrical engineering, from Stanford University, Stanford, Calif. He is currently a Professor at the Department of the Electrical Engineering of the National University of La Plata and a member of its Industrial Electronics, Control, and Instrumentation Laboratory (LEICI). He is also a Member of the Comisión de Investigaciones Cientificas de la Provincia de Buenos Aires. He was a Visiting Professor at Yale University in 1983 and 1994, and at the University of Illinois at Chicago in 1996, 1997, 1999, and 2003. His research interests are in the area of statistical signal and array processing with biomedical, control, and communications applications, and nonlinear control systems.

J. E. Cousseau was born in Mar del Plata, Argentina. He received the B.S. from the Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina, in 1983, the M.S. degree from COPPE/Universidade Federal do Rio de Janeiro (UFRJ), Brazil, in 1989, and the Ph.D. from COPPE/UFRJ, in 1993, all in electrical engineering. Since 1984, he has been with the undergraduate Department of Electrical and Computer Engineering at UNS. He has also been with the graduate program at the same university since 1994. He is a researcher of the National Scientific and Technical Research Council (CONICET) of Argentina. He has been involved in scientific and industrial projects with research groups and companies from Argentina, Brazil, Spain, USA, and Finland. He is coordinator of the Signal Processing and Communication Laboratory (LaPSyC) at UNS. His research interest are related to the areas of FIR and IIR adaptive signal processing, and communication applications.