# Fractionally Spaced Equalization Using CMA: Robustness to Channel Noise and Lack of Disparity

Inbar Fijalkow, Azzèdine Touzni, and John R. Treichler, *Fellow, IEEE*

***Abstract*—In the noise-free case, the fractionally spaced equalization using constant modulus (FSE-CM) criterion has been studied previously. Its minima were shown to achieve perfect equalization when zero-forcing (ZF) conditions are satisfied and to be able to still achieve fair equalization when there is lack of disparity. However, to our best knowledge, the effect of additive channel noise on the FSE-CM cost-function minima has not been studied. In this paper, we show that the noisy FSE-CM cost function is subject to a smoothing effect with respect to the noise-free cost function, the result of which is a tradeoff between achieving zero forcing and noise enhancement. Furthermore, we give an analytical closed-form expression for the loss of performance due to the noise in terms of input-output mean square error (MSE). Under the ZF conditions, the FSE-CM MSE is shown to be mostly due to output noise enhancement and not to residual intersymbol interference (ISI). When there is lack of disparity, an irreducible amount of ISI appears independently of the algorithm. It is the lower equalizability bound for given channel conditions and equalizer length—the so-called minimum MSE (MMSE). The MMSE lower bound is the sum of the MMSE and of additional MSE mostly due to noise enhancement. Finally, we compare the FSE-CM MSE to this lower bound.**

## I. Introduction

COMMUNICATION systems using digital transmission are subject to dispersion due to transmitter filtering, multipath propagation, and receiver filtering, all resulting in intersymbol interference (ISI) [23]. Noise present at the receiver input also perturbs the signal observed at the detector input. At the receiver, the distortions must be compensated in order to retrieve the transmitted symbols. This is referred to as channel equalization.

Most equalization techniques, such as those using the least mean square (LMS) algorithm, use a training sequence to allow the receiver to choose the linear equalizer taps. One obvious drawback of such an approach is that the training causes a reduction of the useful information rate with respect to the total information rate. In other words, training leads to an increase of the necessary bandwidth on the duration required to send a given amount of data. Moreover, there are digital communications systems applications, such as broadcast configurations, where a training scenario is not feasible ([23]). Consequently, *blind equalization* (i.e., without training sequence) has received an increasing interest during the last few years, and, more particularly, the so-called Godard algorithm [12] or constant modulus algorithm (CMA) ([22]).

In digital communications, CMA is mainly used in a fractionally spaced context (since fractional spacing makes it easier to perform other tasks in the receiver system such as synchronization and phase recovery) [23]. Recent work has shown that *channel diversity* produced either by oversampling the received analog signal (i.e., temporal diversity) and/or by using an array of sensors at the receiver (i.e., spatial diversity) may allow *perfect equalization* in a noise-free context (see [20] and [4], for example). However, such perfect equalization depends on the *equalizer length* and the presence of *channel disparity* due to spatio-temporal diversity [20], [4], [19], [18]. These conditions are known as the zero-forcing (ZF) conditions since, if satisfied, they allow the equalizer to force the nonmain symbols to zero. Still, CMA was not studied that way before [17] and [7]. In [9], we have shown that in a noise-free context, the minima of fractionally spaced equalizer using constant modulus (FSE-CM) criterion still equalizes in most cases even if the ZF conditions are not satisfied, whereas the other recently described methods fail. One may claim that ZF conditions are almost always satisfied (i.e., it is very unlikely that subchannels have common zeros). However, we will see that optimal performance in presence of noise depend on how well these conditions are satisfied (i.e., a measure of disparity). Therefore, pure lack of disparity can be seen as an interesting borderline case. To our best knowledge, no analytical study exists under realistic operating conditions, i.e., when additive noise is present, and the channel is possibly affected by a lack of disparity. This is the focus of this paper.

This work is motivated by the desire to understand what is the "best" performance to be expected when a criterion such as the FSE-CM is applied, whether the channel diversity induces enough disparity or not. Such a study will also allow (for given channel disparity conditions and signal-to-noise ratio (SNR)), comparison of the optimal performance of FSE-CM with other recently developed techniques for fractionally spaced equalization. Many equalization criteria result from assuming channel disparity, including both multichannel linear prediction [1], [19], [11] and subspace methods [10], [18]. We will show that not only FSE-CM is more robust to a lack of disparity than other known algorithms; it also has almost optimal asymptotic performances when there is disparity. This property may seem contrary to intuition since FSE-CM is based on fourth-order statistics. To prove this result, we will study the equalizer minimizing the FSE-CM criterion. We start

Manuscript received December 1, 1995; revised August 21, 1996.

I. Fijalkow and A. Touzni are with the Equipe de Traitement des Images et des Signaux (ETIS), Ecole Nationale Supérieure de l'Electronique et de ses Applications (ENSEA), 95014 Cergy-Pontoise Cedex, France (e-mail: fijalkow@ensea.fr; touzni@ensea.fr).

J. R. Treichler is with Applied Signal Technology Inc., Sunnyvale, CA 94086 USA (e-mail: jrt@appsig.com).

Publisher Item Identifier S 1053-587X(97)00517-5.

by fully expressing the FSE-CM cost function in the presence of white additive noise. We will describe the new cost-function minima as a perturbation of the noise-free minima, which is proportional to the SNR, for high SNR. This will enable us to measure the corresponding performance in terms of the input-output mean square error (MSE). To interpret the FSE-CM MSE, we compare it with the lowest achievable input-output MSE, i.e., the so-called minimum MSE (MMSE) or lower equalizability bound. To do so, we need also to express the MMSE in terms of the channel disparity and SNR.

*Organization:* In Section II, the noisy fractionally spaced propagation channel model is presented. In Section III, we establish the FSE-CM cost-function criterion in the presence of white additive noise, and we characterize the minima of the new cost-function. In Section IV, we use the results of the previous sections to study the performance of the FSE-CM minima in terms of the input-output error power both analytically and with simulations.

## II. Fractionally Spaced Channel Model

In this section, we recall the channel model for the fractionally spaced channel, and we define notations to be used in the following sections. Fractionally spaced equalization depends on exploiting temporal and/or spatial diversity at the receiver to recover the transmission signal. The diversity is temporal when the received signal is sampled faster than the required Baud rate and is called spatial when the receiver combines the outputs of an array of sensors. Both schemes have been shown to be equivalent in terms of the equalization problem as they can both be represented by a single input/multiple outputs (SIMO) model (see [18] for example).

### A. Problem Formulation

*1) Channel Model:* The fractionally spaced communication system can be seen to be a SIMO transfer function with a zero-mean single input $s(n)$, the transmission symbol stream, and $L$ outputs corrupted by additive noise (as displayed on Fig. 1). $L$ denotes either the temporal oversampling factor or the number of sensors. One can consider the $L$ functions $c_k(z), k = 1,\dots,L$ to describe the transfer characteristics from the transmitted sequence to either the $k$th sampled received slice in the oversampling case or to the $k$th sensor of the sensor array (see [18] or [7]). We collect them in a $L$-long vector, where each component is a scalar transfer function $\vec{\mathrm{c}}(z) = (\mathrm{c}_1(z),\dots,\mathrm{c}_L(z))^\top$. $\vec{\mathrm{c}}(z) = \sum_p \vec{c}(p) z^{-p}$, where each $\vec{c}(p)$ collects the corresponding $L$ impulse response terms with temporal indice $p$. The additive noise is described by the vector $\vec{w}(n) = (w_1(n),\dots,w_L(n))^\top$. The baseband received signal results in a $L$ variate vector denoted $\vec{r}(n) = (r_1(n),\dots,r_L(n))^\top$.

$$\vec{r}(n) = \sum_p \vec{c}(p) s(n-p) + \vec{w}(n) \tag{1}$$

which can be written as $\vec{r}(n) = [\vec{\mathrm{c}}(z)]s(n) + \vec{w}(n)$, meaning that the signal is the sum of the output of the transfer function $\vec{c}(z)$ driven by $s(n)$ and of the $L$-variate noise $\vec{w}(n)$. We assume, herein, that the channel $\vec{c}(z)$ encompasses the effects of the transmission filter, channel response, reception filter, and any other linear dispersive effect encountered between the transmitter's symbol generation and the receiver's equalizer.

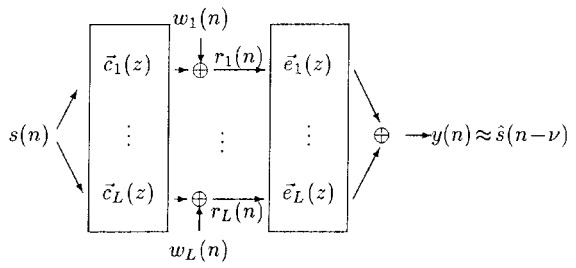



Fig. 1. Noisy fractionally spaced equalization scheme.

The hypotheses on the channel model are as follows:

- H-1) Each $\mathrm{c}_k(z)$ is assumed to be causal with the finite impulse response (FIR) of a degree that is less or equal to $Q$.
- H-2) The noise $\vec{w}(n)$ is a zero-mean, circular (i.e., each $E[w_k^2] = 0$ for complex valued signals), Gaussian, temporally and spatially white ($E[\vec{w}(n)\vec{w}^{*\top}(m)] = E[|w|^2]I$ if $m = n$, 0 otherwise. $I$ stands for the $L \times L$ identity matrix).
- H-3) $s(n)$ is an i.i.d. (i.e., independent and identically distributed) sequence, zero-mean, circular, with variance $E[|s|^2] = 1$ and sub-Gaussian (i.e., $E[|s|^4] - 2E[|s|^2]^2 - |E[s^2]|^2 < 0$).
- H-4) $\vec{w}(n)$ is independent from the input $s(n)$.

The sub-Gaussian assumption in H-3 is not very restrictive since it is satisfied for virtually all digital communication modulations (including PAM, PSK, and QAM). Note that the whiteness and Gaussian assumptions in H-2 are not crucial but simplify the expressions in the sequel.

*2) Equalization:* Solutions of the multivariate channel linear equalization problem hinges on pseudoinverting the multichannel transfer function by choosing an equalizer transfer function $\vec{\mathrm{e}}(z)$ so that $y(n) = [\vec{\mathrm{e}}^\top(z)]\vec{r}(n)$ is a good estimation of $s(n-\nu)$ (up to a complex valued constant) with $\nu$ being the delay of the combined channel/equalizer (see Fig. 1). Since $y(n)$ must be scalar, the equalizer transfer function is an $L$ long vector $\vec{\mathrm{e}}(z) = (\mathrm{e}_1(z),\dots,\mathrm{e}_L(z))^\top$. Each component is an FIR filter with degree $N-1$ so that $\vec{\mathrm{e}}(z) = \sum_{p=0}^{N-1} \vec{e}(p)\, z^{-p}$. The associated equalizer impulse response $\vec{e} = (\vec{e}^\top(0),\dots,\vec{e}^\top(N-1))^\top$ is a $NL$-long vector.

Setting $\vec{R}(n) = (\vec{r}^\top(n), \vec{r}^\top(n-1),\dots,\vec{r}^\top(n-N+1))^\top$, the regressor vector of the $N$ last observations of the $L$-variate $\vec{r}(n)$, we obtain the following linear model:

$$\vec{R}(n) = \mathbf{C}S(n) + \vec{W}(n) \tag{2}$$

where $S(n) = (s(n),\dots,s(n-Q-N+1))^\top$, $\vec{W}(n) = (\vec{w}^\top(n),\dots,\vec{w}^\top(n-N+1))^\top$, and $\mathbf{C}$ is a $NL\times(N+Q)$ block Sylvester convolution matrix [2] defined by the $(Q+1)L$-long impulse response of the channel $\vec{C} = (\vec{c}^\top(0),\dots,\vec{c}^\top(Q))^\top$, as

$$\mathbf{C} = \begin{bmatrix} \vec{c}(0) & \cdots & \vec{c}(Q) & 0 & \cdots & 0 \\ 0 & \vec{c}(0) & \cdots & \vec{c}(Q) & \cdots & 0 \\ & & \ddots & & \ddots & \\ 0 & \cdots & 0 & \vec{c}(0) & \cdots & \vec{c}(Q) \end{bmatrix}.$$

The scalar output $y(n)$ can be written as $y(n) = \vec{c}(n)^\top \vec{R}(n) = h^\top S(n) + \vec{c}^\top \vec{W}(n)$, where $h = \mathbf{C}^\top \vec{c}$ is the $(N+Q)$-length global impulse response of the channel/equalizer system. Note that the equalizer output signal $y(n)$ depends on the additive filtered noise $\vec{c}^\top \vec{W}(n)$. The larger the equalizer taps are, the larger the output noise power is. Such an increase in noise level is called noise enhancement, and its importance will be quantified shortly.

### *B. Noise-Free Equalizability Conditions*

Before analyzing the effect of noise, let us recall some relevant results in the noise-free fractionally spaced equalization case. When the channel is noise-free, the equalizer output is given by $y(n) = h^\top S(n)$ so that the equalization problem turns on solving

$$\mathbf{C}^\top \vec{c} = h_\nu = (0...010...0)^\top \tag{3}$$

where the nonzero entry is the $(\nu+1)$th component. As a practical matter, solving this equation within a scale factor and a complex rotation is equally good. For sake of simplicity, we constrain our study that of (3). In terms of transfer functions, the problem is stated as $\vec{\mathrm{e}}(z)^\top \vec{\mathrm{c}}(z) = z^{-\nu}$, which is referred to as the ZF equalization problem. Recent contributions (see [17] for one of the first studies) have shown that in the noise-free case (and for $L \geq 2$), ZF equalization is achievable under the following assumptions:

- C-1) There are no common roots to all $\mathrm{c}_1(z), \ldots, \mathrm{c}_L(z)$ (the so-called identifiability condition).
- C-2) $N \geq Q$ (the so-called equalizer length condition).

In fact, C-1 is a necessary condition, unlike C-2. More precisely, the condition ensuring that the channel convolution matrix has more rows than columns is $N(L-1) \geq Q$. This condition is indeed not achievable for $L = 1$. For $L = 2$, it is equivalent to C-2. For $L > 2$, it is weaker than C-2, which is then a sufficient condition only. However, since we want to take the same number of taps for each $\mathrm{e}_k(z)$ (we do not know how to select the relevant $k$) and since $Q/(L-1)$ may not be an integer, we assume in the sequel that the equalizer is long enough to ensure that C-2 is satisfied.

Under the ZF conditions, $\mathbf{C}$ is a full column-rank matrix [2], [15]. When satisfied, (3) can be solved, and any combined channel/equalizer impulse response $\vec{c}$ of $(N+Q)$-length is achievable. In particular, the canonical vectors $h_\nu$ that achieve perfect ZF can be obtained. Under the noise-free assumption, the previous conditions allow perfect equalization (i.e., $y(n) = h_\nu^\top S(n) = s(n-\nu)$), which is achievable by many algorithms, among which, there is the FSE-CM algorithm [7]. Moreover, the ZF conditions imply that $\vec{\mathrm{c}}(z)$ is a minimum-phase transfer function [15]. Consequently, second-order statistics (SOS)-based methods can be applied [20], [21].

When the identifiability condition is not satisfied, we assume that all subchannels $\mathrm{c}_k(z)$ have $Z_0$ common zeros (this is referred to as lack of disparity). In this case, identification is not achievable using only SOS, and most of the previously defined methods fail (even in the noise-free case), especially when the common zeros form a nonminimum phase scalar transfer function. Under the equalizer length assumption (i.e., $N \geq Q$), the common roots of all subchannels can be factored, and the convolution matrix can be factored as a product of block-Sylvester matrices as $\mathbf{C} = \underline{\mathbf{C}}\mathbf{C}_0$ (see Appendix A). $\mathbf{C}_0$ is the full row-rank $(N+Q-Z_0)\times(N+Q)$ convolution matrix associated with the common zeros (put together in $\mathrm{c}_0(z)$), and $\underline{\mathbf{C}}$ is the full column-rank $(NL) \times (N+Q-Z_0)$ matrix associated with the remaining part of the multichannel transfer function. Indeed, the corresponding channel equalizer transfer function is given by

$$\mathrm{h}(z) = \mathrm{c}_0(z)\underline{\vec{\mathrm{c}}}(z)^\top \vec{\mathrm{e}}(z) = \mathrm{c}_0(z)\underline{\mathrm{e}}(z)$$

where any scalar degree $N+Q-Z_0$ polynomial $\underline{\mathrm{e}}(z)$ is achievable. In this case, the equalization problem is turned into the nonfractional equalization of a degree $Z_0$ FIR channel $\mathrm{c}_0(z)$ using the degrees of freedom not consumed in achieving the best value of $\underline{\mathrm{e}}(z)$ (see [9]).

Recent studies show that in a noise-free context, the FSE-CM criterion exhibits some robustness with respect to the loss of the disparity condition [9]. One important question remains to this date:

*How robust is the FSE-CM criterion when the received signal is affected by additive noise?*

To answer to this question, we will examine the cases where the ZF conditions are achieved and where there is lack of channel disparity.

### *C. Performance Measure*

While bit error rate would be the natural performance measure for digital transmitted symbol, the relationship between it and the equalizer coefficients is a very complicated one. As a result, it is customary to use the input-output squared error present at the receiver decision circuit instead since in the limit of small amount of Gaussian noise, the mean square input-output error power (i.e., MSE) and the bit error rate are known to be monotonically related. In this light, we will use the MSE.

For a given equalizer $\vec{c}$, the independence assumption H-4 and the whiteness assumption on $S(n)$ and $\vec{W}(n)$ (in H-2 and H-3) yields the corresponding MSE

$$\begin{aligned} \mathrm{MSE} &= E[|y(n) - s(n-\nu)|^2] \\ &= E[|(h-h_\nu)^\top S(n)|^2] + E[|\vec{c}^\top \vec{W}(n)|^2] \\ &= \|h - h_\nu\|^2 + \gamma\|\vec{c}\|^2 \end{aligned} \tag{4}$$

where $\gamma$ denotes the noise-to-signal power ratio $\frac{E[|w|^2]}{E[|s|^2]}$. The MSE is the sum of a ZF measure, i.e., distance between the global channel equalizer impulse response $h$ and the optimal ZF $h_\nu$ and of a noise enhancement measure $\gamma\|\vec{c}\|^2$, i.e., the amount that the received noise is enhanced by the equalizer.

## III. The FSE-CM Criterion

### *A. FSE-CM Algorithm*

Coefficients for a blind equalizer can be attained by minimizing the FSE-CM or Godard cost-function [12], [22], which is written as

$$J(\vec{c}) = E[(r_2 - |y(n)|^2)^2] \tag{5}$$

where $y(n) = \vec{c}(n)^\top \vec{R}(n)$ is the equalizer output $r_2 = \frac{E[|s|^4]}{E[|s|^2]}$. Note that for unit variance input, the dispersion constant $r_2$ is the input kurtosis $\rho = \frac{E[|s|^4]}{E[|s|^2]^2}$. One way to minimize $J(\vec{c})$ is to use the associated stochastic gradient descent algorithm, which is written in the fractionally spaced context as

$$\vec{c}(n+1) = \vec{c}(n) + \mu y(n)(r_2 - |y(n)|^2)\vec{R}^*(n) \qquad (6)$$

where $\mu$ is a small positive step size. this algorithm is called the fractionally spaced equalization by constant modulus algorithm (FSE-CMA). For $L = 1$, it is known as Godard algorithm [12] or CMA [22].

Note that from averaging theory, the extrema of the cost function $J(\vec{c})$ are the possible convergence settings of the algorithm (6). The knowledge of the noisy FSE-CM extrema allows us also to analyze the mean convergence points of the FSE-CM algorithm (6) under noisy observations.

### *B. FSE-CM Noisy Cost Function*

In order to have a better understanding of the equalizer settings minimizing the FSE-CM criterion, we first establish a meaningful expression for the FSE-CM cost function (5) in noisy channel conditions.

*Proposal 1:* Under the hypotheses H-1 and H-4, the FSE-CM cost function $J(\vec{c})$ can be expressed as

$$J(\vec{c}) = J_0(h) + \gamma||\vec{c}||^2\big\{2\big(\rho_g||h||^2 - \rho\big) + \rho_g\,\gamma||\vec{c}||^2\big\}. \quad (7)$$

$J_0(h) = E[(\rho - |h^\top S(n)|^2)^2]$ is the noise-free cost function (studied in [7]). $\rho = \frac{E[|s|^4]}{E[|s|^2]^2}$ is the input signal kurtosis, $\rho_g$ is a Gaussian signal kurtosis (equal to 3 in the case of real-valued signals and to 2 in the complex-valued case) and $\gamma = \frac{E[|w|^2]}{E[|s|^2]}$ is the noise-to-signal power ratio.

According to Proposal 1, one can see that the FSE-CM cost-function minimization in the noisy context is equivalent to the noise-free cost-function $J_0(h)$ minimization (over the vectors $h = \mathbf{C}^\top\vec{c}$) regularized by the additional deterministic factor $\Phi(\vec{c}) = ||\vec{c}||^2\{2(\rho_g||h||^2 - \rho) + \rho_g\gamma||\vec{c}||^2\}$. Intuitively, we may represent this smoothing effect as a quadratic local deformation (with respect to $\gamma$) of the cost-function $J_0(h(\vec{c}))$.

Note that for most input sequences $s(n)$ (like PAM/QAM for instance), $\rho < \rho_g - 1$; therefore, $||h||^2 > \rho/\rho_g$ is true for $h$ not too close to the origin $h = 0$, which is the maximum of $J(\vec{c})$. Then, $\Phi(\vec{c})$ has the desirable property to be a positive convex quartic function of $||\vec{c}||$. The main consequence of the regularization is to forbid the equalizer norm to be too high, reducing the noise enhancement contribution $\vec{c}^\top\vec{W}(n)$ in the equalizer output signal $y(n)$. The smoothing effect is all the more important when the noise power increases. It must be noted that compared with the MSE criterion (4), the constraint on the norm of $\vec{c}$ is much stronger, limiting all the more noise enhancement.

*Proof of Proposal 1:* We denote $\vec{R}(n) = \mathbf{C}S(n) + \vec{W}(n) = \vec{R}_s(n) + \vec{R}_w(n)$ and $y(n) = \vec{c}^\top\vec{R}(n) = y_s(n) + y_w(n)$, where subscripts $s$ and $w$ denote, respectively, the signal and noise contributions. Under the independence hypothesis H-4 and the zero-mean assumption, a straightforward calculus leads to $E[(r_2 - |y|^2)^2] = E[(r_2 - |y_s|^2)^2] - 2r_2E[|y_w|^2] + E[|y_w|^4] + 4E[|y_s|^2]E[|y_w|^2] + E[y_s^2]E[y_w^{*2}] + E[y_s^{*2}]E[y_w^2]$. Under H-2 and H-3, $E[|y_w|^2] = E[|w|^2]||\vec{c}||^2 = \gamma||\vec{c}||^2$, and $E[|y_s|^2|y_w|^2] = E[|s|^2]E[|w|^2]\ ||\vec{c}||^2||h||^2 = \gamma||\vec{c}||^2||h||^2$. In the complex case, $E[y_s^2]E[y_w^{*2}] = E[y_s^{*2}]E[y_w^2] = 0$ by circularity. In the real case, $E[y_s^2]E[y_w^{*2}] = E[y_s^{*2}]E[y_w^2] = \gamma||\vec{c}||^2||h||^2$. Because the noise is Gaussian, one can express $E[|y_w|^4]$ with second-order moments only, as $E[|y_w|^4] = 2E[|y_w|^2] + |E[y_w^2]|^2$. In the complex case, $E[|y_w|^4] = 2E[|w|^2]^2||\vec{c}||^4 = 2\gamma^2||\vec{c}||^4$ by circularity. In the real case, $E[|y_w|^4] = 3E[|w|^2]^2||\vec{c}||^4 = 3\gamma^2||\vec{c}||^4$. In both cases, $E[|y_w|^4] = \rho_g\gamma^2||\vec{c}||^4$. Adding all terms, we get (7). Note that (7) for the noisy FSE-CM cost function can easily be extended to the case of non-Gaussian or spatially colored noise. □

### *C. FSE-CM Cost-Function Minimization*

Recalling previous results in the noise-free case [7], we want to compare them to the noisy cost-function extrema to measure the loss of performances induced by the noise. Since the noisy FSE-CM cost-function is continuous in terms of $\gamma$, we assume by an homotopic continuity argument that the extrema of the noisy case can be deduced for small $\gamma$ by a slight perturbation in terms of $\gamma$ of the noise-free extrema. Small $\gamma$ (large SNR) values correspond to one transmitter and a not-too-great amount of additive noise. Large $\gamma$ (negative SNR) values correspond to the presence of an interferer and/or too a high amount of noise. In this last case, the main term of the cost function is $\gamma^2||\vec{c}||^4$, the minimal value of which is given by $\vec{c} = 0$, which is the value maximizing $J_0$. The previous continuity assumption does not hold any longer and is not the focus of our study.

For any $\gamma$, the search for an analytical expression of the FSE-CM cost-function minima may be addressed only after noting that the cost-function $J(\vec{c})$ (7) depends explicitly on both $\vec{c}$ and $h$ and, therefore, that it cannot be reduced to a function of $h$ only, as in the noise-free case. Because of the constraint on $\vec{c}$, the cost-function extrema form a finite subset of equalizer settings. Whereas in the noise-free case ($J_0$ depends on $h$ only), each extremum in terms of $h$ corresponds to a $NL - (N+Q)$-dimensional dense subspace of equalizer settings.

Therefore, we propose a two-step procedure minimization. First, we minimize $J(\vec{c})$ over the subspace of vectors $\vec{c}$ under the constraint $h = \mathbf{C}^\top\vec{c}$ (for a given $h$). The resulting value of $\vec{c}$ is a function of $h$, which is denoted $\vec{c}(h)$. Then, the minimization of $J(\vec{c}(h))$ is performed over the subspace of $(N+Q)$-long vectors $h$. Invoking (7), the minimization is equivalent to

$$\min_{\vec{c}} J(\vec{c}) = \min_{h}\Big\{J_0(h) + \gamma \min_{\vec{c},h=\mathbf{C}^\top\vec{c}}\Phi(\vec{c})\Big\}. \qquad (8)$$

Note that the procedure is simplified because $J_0$ is a function of $h$ only. The first step also consists of the smoothing cost-function $\Phi(\vec{c})$ minimization only. The quartic cost-function $\Phi(\vec{c})$ minimization under the linear constraint $h = \mathbf{C}^\top\vec{c}$ can be performed using Lagrange multiplier technique. The calculus depends on solving a second-order equation in terms of $||\vec{c}||^2$ (see Appendix B), the analytical solution for which is too

complicated to help understanding the effect of noise. In order to find an analytical and meaningful expression for the minima, we focus on the case when $\gamma$ is small.

For small $\gamma$ (high SNR), $\Phi(\vec{c})$ can be approximated by $2\big(\rho_g\|h\|^2-\rho\big)\ \|\vec{c}\|^2+o(1)$, which is the simplest quadratic function of $\vec{c}$, so that the calculus can easily be performed. In the sequel, we use this approximation to evaluate an analytical closed-form expression of the FSE-CM criterion under ZF and lack of channel disparity conditions.

*1) Minima Under ZF Conditions:* Under the ZF condition, all vectors $h$ are achievable so that the first step minimization results in a first-order approximation of $\vec{c}(h)$, with respect to $\gamma$, and is given by

$$\vec{c}(h)=\mathbf{C}^*(\mathbf{C}^\top\mathbf{C}^*)^{-1}h+o(1) \tag{9}$$

where $*$ denotes the conjugation. The proof of (9) appears in Appendix B.

Next, we are looking at the second step of the minimization procedure, i.e., the minimization of $J(\vec{c}(h))$ versus $h$.

$$J(\vec{c}(h))=J_0(h)+2\gamma(\rho_g\|h\|^2-\rho)\ h^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h^*+o(\gamma)$$

where $J_0(h)=E[(\rho-|h^\top S(n)|^2)^2]$ denotes the FSE-CM noise-free normalized cost function. The cost-function extrema are the points for which the cost-function derivative[1] with respect to $h$ equals 0.

$$\begin{aligned}\frac{1}{2}\vec{\nabla}_h\{J(\vec{c}(h))\}=\Delta(h)h^*+\gamma(\rho_g h^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h^*h^*\\ +(\rho_g\|h\|^2-\rho)(\mathbf{C}^\top\mathbf{C}^*)^{-1}h^*)+o(\gamma).\end{aligned} \tag{10}$$

$\Delta(h)=(\rho_g\|h\|^2-\rho)I-(\rho_g-\rho)\mathrm{diag}(hh^{*\top})$ with $\mathrm{diag}(A)$ is defined as the matrix extracted from $A$ with the same diagonal entries and 0 elsewhere. $2\Delta(h)h^*$ is the noise-free cost-function derivative as calculated in [7] in the case of real signals.

Under the ZF conditions and noise-free case, the FSE-CM $J_0(h)$ extrema points ($\Delta(h)h^*=0$) are classified as follows (see [7]):

- maximum for $h=0$
- global minima for $h=h_\nu=(0...010...0)^\top$
- saddle points. All nonzero components of $h$ are equal to $\kappa_M=\sqrt{\rho/(\rho_g(M-1)+\rho)}$ up to a modulus 1 factor, where $M$ is the number of nonzero elements of $h$. The solutions of $\Delta(h)h^*=0$ are by definition of the cost-function insensitive to any phase rotation $e^{j\theta}$.

We want to extend this result to the noisy case, but unfortunately, finding an exact analytic solution of (10) in the noisy case is difficult. That is why, in the following, we will look for a first-order solution to (10) with respect to $\gamma$ by supposing $\gamma$ small enough. The solution points of (10) are then assumed to be a perturbation in terms of $\gamma$ of the noise-free equation $\Delta(h)h^*=0$. Assuming a small perturbation of $h_\nu$ in terms of $\gamma$, our task is to see how much the new stationary points are deviated in noisy case.

[1] $f(z), z=x+jy$ being a complex function, we consider the usual complex derivative as $\partial f/\partial z=\partial f/\partial x-j\partial f/\partial y$ and $\partial f/\partial z^*=\partial f/\partial x-j\partial f/\partial y$ so that the gradient of a cost function is defined by $\vec{\nabla}_h\{J\}=(\partial J/\partial h_1^*,\ldots,\partial J/\partial h_{N+Q}^*)^\top$.

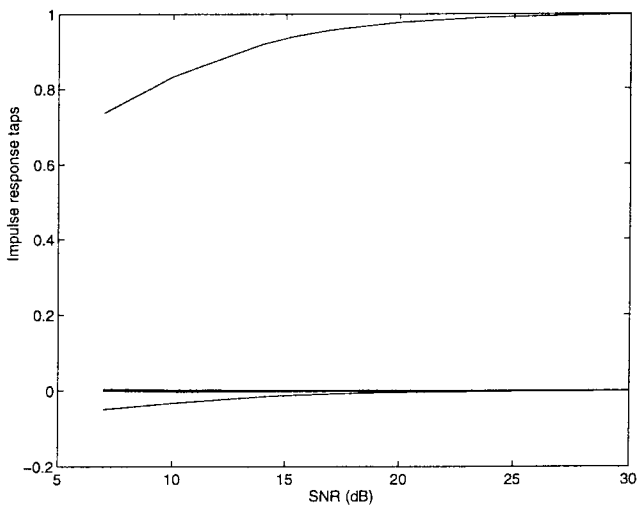



Fig. 2. Under ZF conditions: Impact of SNR on $h$, the global channel/equalizer minimizing FSE-CM criterion. $L=2$, $Q=2$, and $N=2$. $h$ is averaged over20 000 iterations of the FSE-CMA at steady state with step-size 0.0001.

*Proposal 2:* We assume that the global channel/equalizer impulse response setting $h_\gamma$ corresponding to a stable solution of $\vec{\nabla}_h\{J(\vec{c}(h))\}=0$ (10) admits an approximation in terms of $\gamma$ as

$$h_\gamma=h_\nu+\gamma\,\bar{h}_\nu+o(\gamma). \tag{11}$$

In this case, $\bar{h}_\nu$ is given by

$$\bar{h}_\nu=-\frac{(2\rho_g-3\rho)}{2\rho}h_\nu^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\nu\ h_\nu-\big(\mathbf{C}^\top\mathbf{C}^*\big)^{-1}h_\nu. \tag{12}$$

This approximation depends only on the expression suggested in (11). In particular, this first-order development does not use the perturbation expression $\bar{h}_\nu$. In fact, the simulations (see Figs. 2 and 3) have shown a good accuracy between the experimental results and the analytical expression so that we have not introduced a second-order approximation in terms of $\gamma$.

*Proof of Proposal 2:* The proof of Proposal 2 comes from a first-order approximation of (10) in terms of $\gamma$. Assuming that the solution of (10) is of the form (11), we have $\Delta(h_\gamma)=\Delta(h_\nu)+$

$$\gamma\big(\rho_g h_\nu^\top\big(\bar{h}_\nu^*+\bar{h}_\nu\big)I-(\rho_g-\rho)\mathrm{diag}\big(h_\nu^\top\ \big(\bar{h}_\nu^*+\bar{h}_\nu\big)\big)\big)+o(\gamma).$$

Since $\Delta(h_\nu)h_\nu^*=0$, and $\mathrm{diag}(\bar{h}^*h_\nu^\top)\,h_\nu^*=\mathrm{diag}(h_\nu\bar{h}^{*\top})\,h_\nu^*=h_\nu h_\nu^{*\top}\,\bar{h}^*$ for any $\bar{h}$, what remains is

$$\begin{aligned}&\Delta(h_\gamma)h_\gamma^*+\gamma\big(\rho_g h_\gamma^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\gamma^*h_\gamma^*\\ &\quad+(\rho_g\|h_\gamma\|^2-\rho)(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\gamma^*\big)+o(\gamma)\\ &\quad=\gamma\Delta(h_\nu)\bar{h}_\nu^*+\gamma\rho\big(h_\nu h_\nu^{*\top}\bar{h}_\nu^*+h_\nu^*h_\nu^\top\bar{h}_\nu\big)\\ &\quad+\gamma\rho_g h_\nu^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\nu^*h_\nu^*+\gamma(\rho_g-\rho)(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\nu^*+o(\gamma).\end{aligned}$$

Adding together this equation and its conjugate, we get $\gamma(\rho_g h_\nu^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\nu\ I+(\rho_g-\rho)(\mathbf{C}^\top\mathbf{C}^*)^{-1})h_\nu+\gamma\Psi_\nu(\bar{h}_\nu^*+\bar{h}_\nu)+o(\gamma)=0$, with $\Psi_\nu$ a diagonal matrix with entry $(\rho_g-\rho)$ when $i\neq\nu+1$ and $2\rho$ when $i=\nu+1$. Note that when $h_\nu$ is real valued, so is $\bar{h}_\nu$. When there is a phase rotation factor

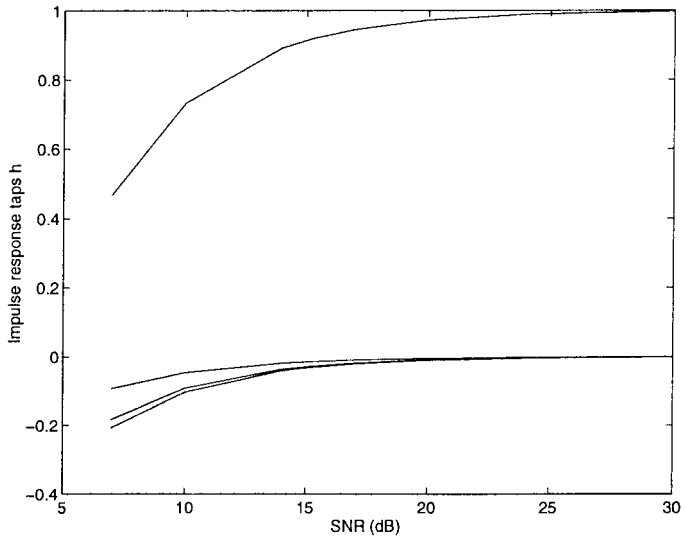



Fig. 3. Under ZF conditions: Impact of SNR on the analytical model of the global channel/equalizer impulse response $h$ using the same channel and equalizer length as in Fig. 2.

on $h_\nu$, a corresponding one appears on $\bar{h}_\nu$. Since it does not affect the impact of $\gamma$ on the $h_\gamma$, we consider the case of real-valued $h_\nu$.

Finally, we can solve the previous system by multipling both sides by $\Psi_\nu^{-1}$ (which is well defined since $\rho \neq \rho_g$). The $\nu+1$ tap of $\bar{h}_\nu$ is equal to $-\frac{2\rho_g-\rho}{2\rho} h_\nu^\top (\mathbf{C}^\top \mathbf{C}^*)^{-1} h_\nu$, the other taps $(i \neq \nu+1)$ are equal to $-\bar{h}_{i-1}^\top (\mathbf{C}^\top \mathbf{C}^*)^{-1} h_\nu$, which can be collected as in (12). □

The justification of (11) comes from an assumption of homotopic continuity of a stable setting when $\gamma$ is small enough. Moreover, we assume that the stability is not changed for $\gamma$ small enough. We must note that the approximation in (11) is done in terms of the global channel/equalizer impulse response, but it is not true for the equalizers settings $\vec{c}$ because of their nonuniqueness in the noise-free case, i.e., $h_\nu = \mathbf{C}^\top \vec{c}$. According to Proposal 2, we are thus able to understand some channel noise effects. Intuitively, $J(\vec{c})$ minima are looked for as a continuous distortion of $J_0$ minima location with respect of $\gamma$. The range of SNR values for which this approximation is valid is checked by means of simulations (see Figs. 2 and 3). In particular, noise introduces two main perturbations on $h_\nu$. First, an attenuation on the main direction given by $h_\nu$ (with respect to the noise-free equalizer norm). If it is not too large, this perturbation does not induce major damage in terms of the ZF conditions. Then, on the other taps, a nonzero contribution induces some residual ISI. Indeed, the approximation (11) is all the more valid when $\bar{h}_\nu$ is small. $\bar{h}_\nu$ depends on $(\mathbf{C}^\top \mathbf{C}^*)^{-1}$, the determinant of which is inversely proportional to the square of distances between different subchannels zeros (see [6]). The channel transfer function is also of great importance to quantify the perturbation, even when the channel identifiability condition is met.

For a given $\gamma$ (assumed small) and a large enough number of equalizer taps, we suggest the following asymptotical (in terms of $N$) approximation of (12):

$$h_\gamma \sim h_\nu - \gamma (\mathbf{C}^\top \mathbf{C}^*)^{-1} h_\nu + o(\gamma). \tag{13}$$

This result is given by neglecting the difference between the term corresponding to $(\nu+1)$th coefficient and the others. The point of interest here is that the resulting expression (13) is close to the first-order approximation of the minimization of the input-output MMSE $E[|y(n) - s(n-\nu)|^2]$ (see [8] for exact expression of the real case).

*2) Simulations:* The simulations were computed in the real-valued case with a 2-D multichannel transfer function $\vec{c}(z)$, where the zeros of $c_1(z)$ and $c_2(z)$ are, respectively, $(-1.4, 0.6)$ and $(1.1, -0.4)$ (from now on, it will be referred to as channel a)). The input signal $s(n)$ is a BPSK sequence, taking the values $\pm 1$ so that $\rho_g = 3$. The equalizer length is set to $N = 2$. In order to validate Proposal 2, the minimum of the FSE-CM cost function in terms of the global impulse response taps $h_\gamma$ are displayed versus SNR (SNR $= 10\log_{10}(\gamma)$ dB) on Fig. 2. In order to minimize the FSE-CM criterion, we run the algorithm in (6) using a step-size $\mu = 10^{-4}$ and a center-spike initial value. The taps are an average value computed at steady state over 20 000 iterations. The results are compared with the analytical model (12) displayed on Fig. 3. The simulation shows the accuracy of expression (12) down to SNR values as low as 10 dB. Similar results were found with larger values of $N$.

*3) Remark:* If we analyze, in detail, both curves, we may notice that for a low SNR, the analytical solution (Fig. 3) leads to a slight overestimation of the noise perturbation. Similar observation was made on many other simulations. It may explained by the fact that the analytical solution is a first-order approximation in terms of the noise. Yet, we have seen previously that when the noise increases, the smoothing effect increases as well. Indeed, a second-order analytical approximation is required to give more accurate results.

*4) Minima Under Lack of Disparity:* When there is loss of channel disparity, $\mathbf{C}^\top \mathbf{C}^*$ is no longer invertible. In fact, the achievable values of $h$ span the range of $\mathbf{C}_0^\top$ so that a minimization over $h$ must be constrained to that subspace. To do so, the following parametrization is proposed: $h = h(d) = \mathbf{C}_0^\top d$, where $d$ is the $(N + Q - Z_0)$-long vector such that $d = \underline{\mathbf{C}}\, \vec{c}$. For a given achievable $h$, there exists a unique $d$ and a subspace of dimension $NL - (N + Q - Z_0)$ of possible settings for $\vec{c}$. In this case, a two-step minimization will also be done versus the unconstrained parameter of the problem $d$.

The first-step minimization of $\Phi(\vec{c}(d))$ (for a given $d$) is similar to (9), where $\mathbf{C}$ is replaced by the full column-rank $\underline{\mathbf{C}}$ and $h$ by $d$

$$\vec{c}(d) = \underline{\mathbf{C}}^* (\underline{\mathbf{C}}^\top \underline{\mathbf{C}}^*)^{-1} d + o(1).$$

The proof appears in Appendix B.

The second step of the minimization procedure consists of minimizing with respect to $d$

$$J(\vec{c}(d)) = J_0(h(d)) + 2\gamma \, ||\vec{c}(d)||^2 (\rho_g ||h(d)||^2 - \rho) + o(\gamma) \tag{14}$$

where the noise-free cost function is $J_0(h(d)) = E[(\rho - |d^\top \mathbf{C}_0 S(n)|^2)^2]$. The derivative of (14) with respect to $d$ leads to

$$\frac{1}{2} \vec{\nabla}_d J(\vec{c}(d)) = \mathbf{C}_0 \Delta\big(\mathbf{C}_0^\top d\big) \mathbf{C}_0^{*\top} d^*$$

$$+\gamma\rho_g d^\top(\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}d^*\mathbf{C}_0\mathbf{C}_0^{*\top}d^* + \gamma\big(\rho_g d^\top\mathbf{C}\mathbf{C}_0^{*\top}d^*-\rho\big)(\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}d^*+o(\gamma) \tag{15}$$

In the noise-free case, the cost-function extrema of $J(\vec{c}(d))$ ($\mathbf{C}_0\Delta(\mathbf{C}_0^\top d)\mathbf{C}_0^{*\top}d^*=0$) correspond to the minimization of the *nonfractional* CM cost function associated with the common zero transfer function $\mathrm{c}_0(z)$ [9]. The extrema can be classified by the following:

- one maximum ($d=0$)
- global minima (when $\mathbf{C}_0^\top d = h_\nu$ is achievable with $d\neq 0$) or saddle points ($\Delta(\mathbf{C}_0^\top d)\ \mathbf{C}_0^{*\top}d^*=0$ and $\mathbf{C}_0^\top d\neq h_\nu$)
- local minima ($\mathbf{C}_0\Delta(\mathbf{C}_0^\top d)\ \mathbf{C}_0^{*\top}d^*=0$ and $d$ does not belong to the previous categories).

In fact, the existence of a noise-free global minimum implies that the corresponding $d$ equals $(\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*h_\nu$, which implies that $h=\mathbf{\Pi_0}h_\nu$, where $\mathbf{\Pi}_0=\mathbf{C}_0^\top(\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*$ is the projection on the range of $\mathbf{C}_0^\top$. Therefore, $h\neq h_\nu$, unless $h_\nu$ is an eigenvector of $\mathbf{\Pi}_0$. In fact, since $\mathbf{C}_0$ cannot be square (it is a $(N+Q-Z_0)\times(N+Q)$ matrix), there should exist no global minima such as $h=h_\nu$ (otherwise, there would be some FIR $\mathrm{d}(z)$ such as $\mathrm{c}_0(z)\mathrm{d}(z)=z^{-\nu}$). However, when $N$ becomes "very large," $\mathbf{C}_0$ tends to become square so that $\mathbf{\Pi_0}$ becomes close to the identity matrix. Of course, as in the nonfractional case, undesired settings may exist as in [3]. However, the larger $N$ is getting, the closer the corresponding channel/equalizer is getting to some $h_\nu$ (see [14]).

Our task, as under the ZF conditions, is to provide an expression of the minima as a small perturbation in terms of $\gamma$. Since we do not know how to make explicit the noise-free minima expression, we focus on the perturbation of these given by $d=d_\nu=(\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*\ h_\nu$. It is motivated by the fact that for a "large enough" equalizer length, it will be a good approximation of the noise-free minima. In order to perform this approximation, we suggest that the approximation error is "smaller" than the perturbation due to the noise. This should hold for "large" values of $N$ and "not too small" values of $\gamma$. Still $\gamma$ must be small enough to allow a first-order approximation in terms of $\gamma$. The validity of these assumptions is checked by simulations in the sequel.

*Proposal 3:* For a small enough $\gamma$, we assume the global channel/equalizer setting $d_\gamma$ is a first-order perturbation of $d_\nu=(\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*\ h_\nu$ in terms of $\gamma$ as

$$d_\gamma = d_\nu+\gamma\bar{d}_\nu+o(\gamma) \tag{16}$$

$$\bar{d}_\nu \approx -\rho_g d_\nu^\top(\underline{\mathbf{C}}^{*\top}\underline{\mathbf{C}})^{-1}d_\nu^*\left(\mathbf{C}_0^*\Psi_\nu\mathbf{C}_0^\top\right)^{-1}\mathbf{C}_0^*\mathbf{C}_0^\top d_\nu - (\rho_g-\rho)\left(\mathbf{C}_0^*\Psi_\nu\mathbf{C}_0^\top\right)^{-1}(\underline{\mathbf{C}}^{*\top}\underline{\mathbf{C}})^{-1}d_\nu. \tag{17}$$

The corresponding global channel/equalizer settings $h_\gamma=\mathbf{C}_0^\top d_\gamma$ can be viewed as a perturbation of $h_\nu$ as

$$h_\gamma\approx h_\nu-\gamma\mathbf{C}_0^\top\left(\mathbf{C}_0^*\Psi_\nu\mathbf{C}_0^\top\right)^{-1}\mathbf{C}_0^* \times\left\{\rho_g d_\nu^\top\left(\underline{\mathbf{C}}^{*\top}\underline{\mathbf{C}}\right)^{-1}d_\nu^*h_\nu+\left(\rho_g-\rho\right)\mathbf{C}_0^\top\left(\mathbf{C}_0^*\mathbf{C}_0^\top\right)^{-1}\times(\underline{\mathbf{C}}^{*\top}\underline{\mathbf{C}})^{-1}\left(\mathbf{C}_0^*\mathbf{C}_0^\top\right)^{-1}\mathbf{C}_0^*h_\nu\right\}+o(\gamma)$$

The symbol $\approx$ in (17) stands for the approximation of $\mathbf{C}_0^\top d_\nu\approx h_\nu$ for a large value of $N$.

*Proof of Proposal 3:* Introducing assumption (16) in (15), we want to do a similar calculation as in the ZF case. Setting $h=\mathbf{C}_0^\top d_\nu=\mathbf{\Pi_0}h_\nu$,

$$\begin{aligned}\Delta\big(\mathbf{C}_0^\top d_\gamma\big)&=\big(\rho_g\big\|\mathbf{C}_0^\top d_\gamma\big\|^2-\rho\big)I-(\rho_g-\rho)\mathrm{diag}\big(\mathbf{C}_0^\top d_\gamma d_\gamma^{*\top}\mathbf{C}_0^*\big)\\&=\Delta\big(\mathbf{C}_0^\top d_\nu\big)+\gamma\big(\rho_g\big(h^\top\mathbf{C}_0^{*\top}\bar{d}_\nu^*+h^{*\top}\mathbf{C}_0^\top\bar{d}_\nu\big)I\\&\quad-(\rho_g-\rho)\mathrm{diag}\big(h^*\bar{d}_\nu^\top\mathbf{C}_0+h\bar{d}_\nu^{*\top}\mathbf{C}_0^*\big)\big)+o(\gamma).\end{aligned}$$

Note that $\mathbf{C}_0^*h=\mathbf{C}_0^*h_\nu$; therefore, $\mathbf{C}_0\Delta(\mathbf{C}_0^\top d_\gamma)\mathbf{C}_0^{*\top}d_\gamma^*=\mathbf{C}_0\Delta(h)h^*+\gamma\mathbf{C}_0\Delta(h)\mathbf{C}_0^{*\top}\bar{d}_\nu^*+\gamma\mathbf{C}_0\rho_g(h_\nu^\top\mathbf{C}_0^{*\top}\bar{d}_\nu^*+h_\nu^{*\top}\mathbf{C}_0^\top\bar{d}_\nu)h^*-\gamma(\rho_g-\rho)\mathrm{diag}(h^*\bar{d}_\nu^\top\mathbf{C}_0+h\bar{d}_\nu^{*\top}\mathbf{C}_0^*))h+o(\gamma)$.

For a large enough $N$, we assume $\mathbf{C}_0^\top d_\nu\approx h_\nu$; therefore, $\mathbf{C}_0\Delta(\mathbf{C}_0^\top d_\gamma)\mathbf{C}_0^{*\top}d_\gamma^*\approx\mathbf{C}_0\Delta(h_\nu)h_\nu^*+\gamma\mathbf{C}_0\Delta(h_\nu)\mathbf{C}_0^{*\top}\bar{d}_\nu^*+\gamma\mathbf{C}_0(\rho_g(\bar{d}_\nu^{*\top}\mathbf{C}_0^*+\bar{d}_\nu^\top\mathbf{C}_0)h_\nu I-(\rho_g-\rho)\mathrm{diag}(h_\nu(\bar{d}_\nu^{*\top}\mathbf{C}_0^*+\bar{d}_\nu^\top\mathbf{C}_0))h_\nu+o(\gamma)$

$$\begin{aligned}&=0+\gamma(\rho_g-\rho)\mathbf{C}_0\big(I-h_\nu h_\nu^\top\big)\mathbf{C}_0^{*\top}\big(\bar{d}_\nu^*+\bar{d}_\nu\big)\\&\quad+\gamma\rho\mathbf{C}_0h_\nu h_\nu^\top\mathbf{C}_0^{*\top}\big(\bar{d}_\nu^*+\bar{d}_\nu\big)+o(\gamma)\\&=\gamma(\rho_g-\rho)\mathbf{C}_0\mathbf{C}_0^{*\top}\big(\bar{d}_\nu^*+\bar{d}_\nu\big)\\&\quad-\gamma(\rho_g-3\rho)\mathbf{C}_0h_\nu\ h_\nu^\top\mathbf{C}_0^{*\top}\big(\bar{d}_\nu^*+\bar{d}_\nu\big)+o(\gamma)\\&=\mathbf{C}_0\Psi_\nu\mathbf{C}_0^{*\top}\big(\bar{d}_\nu^*+\bar{d}_\nu\big)+o(\gamma).\end{aligned}$$

Thus, $\bar{d}_\nu$ also satisfies $\mathbf{C}_0\Psi_\nu\mathbf{C}_0^{*\top}\bar{d}_\nu^*=$

$$-\big\{\rho_g d_\nu^\top(\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}d_\nu^*\ \mathbf{C}_0h_\nu+(\rho_g-\rho)\ (\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}d_\nu^*\big\}+o(1).$$

Since $\Psi_\nu$ is a diagonal positive matrix, $\mathbf{C}_0\Psi_\nu\mathbf{C}_0^{*\top}$ is invertible, $\bar{d}_\nu^*=-\rho_g d_\nu^\top(\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}d_\nu^*\ (\mathbf{C}_0\Psi_\nu\mathbf{C}_0^{*\top})^{-1}\ \mathbf{C}_0\mathbf{C}_0^{*\top}d_\nu^*-(\rho_g-\rho)\ (\mathbf{C}_0\Psi_\nu\mathbf{C}_0^{*\top})^{-1}(\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}d_\nu^*+o(1)$, which immediately yields (17). □

*5) Simulations:* The 2-D ($L=2$) real multichannel vector $\vec{c}(z)$ is defined by the zeros of each transfer function as $c_1(z)=(-1.4,-0.4)$ and $c_2(z)=(1.1,-0.4)$ (from now on, this is referred to as channel b)). The observation number is set to $N=8$. Because we do not want to deal with eventual local minima when running the algorithm, it was initialized very close to a global minima setting.

Fig. 4 displays the impulse response taps of $h_\gamma$ (which are obtained by running the algorithm to minimize the criterion as in the previous simulation) versus SNR. Note that $h_\gamma$ is very close to a canonical vector for a high enough SNR. In Fig. 5, we display the analytical impulse response model introduced in Proposal 3. We can see that both curves are very close.

## IV. FSE-CM Performances: MSE

In this section, we are interested in the FSE-CM criterion in terms of equalizability, i.e., a measure of the FSE-CM performance in equalizing a given channel in the presence of additive noise. Furthermore, we want to compare the FSE-CM MSE with an analytical expression of the lower achievable minimun MSE (MMSE) over the class of linear $NL$-long equalizers, i.e., $\min_{\vec{c},\nu}\ E[|y(n)-s(n-\nu)|^2]$. For a given delay $\nu$, the corresponding equalizer $\vec{c}$ is selected by minimization of the MSE (see [8]). The optimal delay $\nu$ is chosen to minimize the resulting MSE. Since the MMSE provides a *lower equalizability bound* [8] for a given SNR, this comparison will provide a measure of the FSE-CM criterion performance.

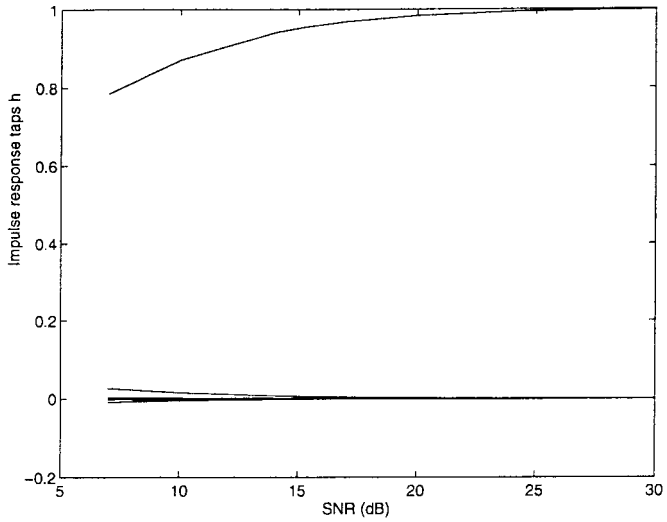



Fig. 4. Lack of disparity: Impact of SNR on $h$, the global channel/equalizer minimizing FSE-CM criterion. $L = 2, Q = 2, Z0 = 1$, and $N = 8$. $h$ is the FSE-CMA average value at steady state.

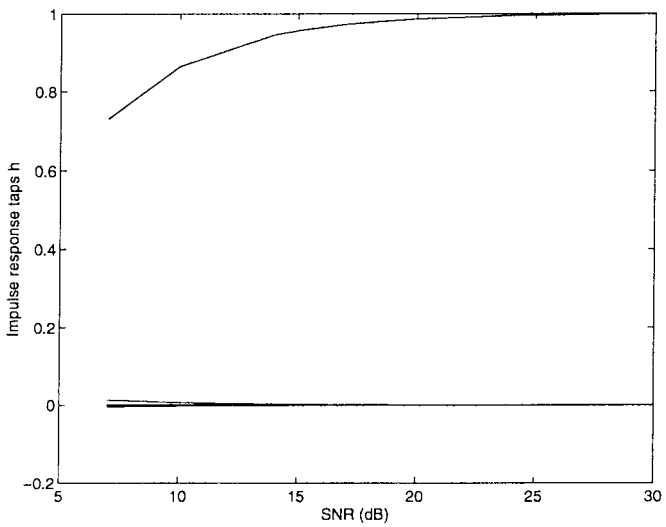



Fig. 5. Lack of disparity: Impact of SNR on the analytical model of the global channel/equalizer impulse response $h$ using the same channel as in Fig. 4.

Note that we do not intend to evaluate the FSE-CM Algorithm performance when the algorithm is used adaptively but only the performance of the equalizer setting minimizing the cost function. Thus, if we want to estimate the global MSE including the algorithm effect related to the step size, we must add to the criterion minima MSE the measure of the stochastic jitter (called excess mean square error (EMSE)) around the mean solution (see [7] for a closed-form expression in the noise-free case).

### A. FSE-CM MSE

Using the results of Proposal 2–3, we can express the FSE-CM MSE for small $\gamma$. To do so, we introduce the FSE-CM minima expressions derived previously in the MSE expression (4). We recall the FSE-CM minima equalizers settings expressions:

- $\mathbf{C}^*(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\nu + o(1)$ under ZF conditions
- $\underline{\mathbf{C}}^*(\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}(\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*h_\nu + o(1)$ under C-2 and a long enough equalizer.

In both cases, $\vec{e} = \vec{e}_\nu + o(1)$; therefore, $h = h_\gamma = h_\nu + \gamma\bar{h}_\nu + o(\gamma)$, resulting in

$$\text{MSE } = \gamma\|\vec{e}_\nu\|^2 + o(\gamma). \tag{18}$$

Recalling the MSE expression (4), $\|h_\gamma - h_\nu\|^2$ is a measure of distance to zero forcing and can, at first (i.e., regarding $\gamma$) be neglected with respect to the second term representing noise enhancement $\gamma\ \|\vec{e}\|^2$. In order to have a more accurate approximation of the MSE (the term depending on $\gamma^2$), one more term is needed in the approximation done in the first minimization step (see Appendix B).

$\|h_\gamma - h_\nu\|^2$ evaluates the residual ISI by an $L^2$ norm, whereas the residual ISI is usually looked at by the following $L^1$ metric:

$$\mu_{\text{ISI}}(h_\gamma) = \frac{\sum_i |h_{\gamma_i}| - \max_i |h_{\gamma_i}|}{\max_i |h_{\gamma_i}|}.$$

$\mu_{\text{ISI}}$ is a measure of the undesired contributions (due to residual ISI) over the desired contribution, represented by the absolutely largest term of the impulse response $h_\gamma$. It is a worst possible case. Conversely, however, $\mu_{\text{ISI}}$ overestimates the problem as viewed by the detector, which, for high SNR and low ISI, views the world as a squared error problem. Therefore, we will consider $\|h_\gamma - h_\nu\|^2$ to be an accurate ZF error power measure.

The other term involved in the MSE is the noise power enhancement by the equalizer $\gamma\|\vec{e}_\gamma\|^2$. The output noise enhancement is expressed as the equalizer output noise to signal ratio

$$\Gamma_\gamma = \frac{E\left[\left|\vec{e}_\gamma^\top\vec{w}(n)\right|^2\right]}{E\left[\left|h_\gamma^\top s(n)\right|^2\right]} = \gamma\frac{\|\vec{e}_\nu\|^2}{\|h_\nu\|^2} + o(\gamma).$$

The expressions of the FSE-CM MSE will be derived in the two cases afterwards. At first, we are expressing the MMSE values for a given channel and SNR in order to compare the FSE-CM MSE to the MMSE.

### B. MMSE: Lower Equalizability Bound

The noisy MSE cost function in (4) is the noise-free MSE cost function $\|h - h_\nu\|^2$ regularized by the simplex quadratic function $\gamma\|\vec{e}\|^2$. Therefore, the MSE cost-function minimization under the constraint $h = \mathbf{C}^\top\vec{e}$ is performed in [8]. It results in the following MMSE expressions:

- $\gamma h_\nu^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\nu + o(\gamma)$, under C-1 and C-2.
- $\|(I - \mathbf{\Pi}_0)h_\nu\|^2 + \gamma h_\nu^\top\mathbf{C}_0^{-L*\top}(\underline{\mathbf{C}}^{*\top}\underline{\mathbf{C}})^{-1}\mathbf{C}_0^{-L}h_\nu + o(\gamma)$ under C-2 only, with $\mathbf{C}_0^{-L} = (\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*$ being $\mathbf{C}_0$ left-inverse.

$\|(I - \mathbf{\Pi}_0)h_\nu\|^2$ is the irreducible ISI due to the channel lack of disparity and to the equalizer being too short and, therefore, failing the conditions for noise-free ZF.

It appears that the terms proportional to $\gamma$ are the same as for FSE-CM in (18), at least in the ZF case. This is not in contradiction with the previous results concerning the FSE-

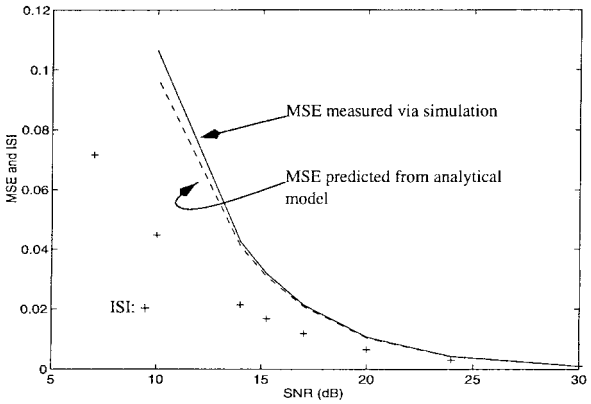



Fig. 6. Under ZF conditions (same simulation conditions as in Fig. 2): Impact of SNR on FSE-CMA MSE and ISI.

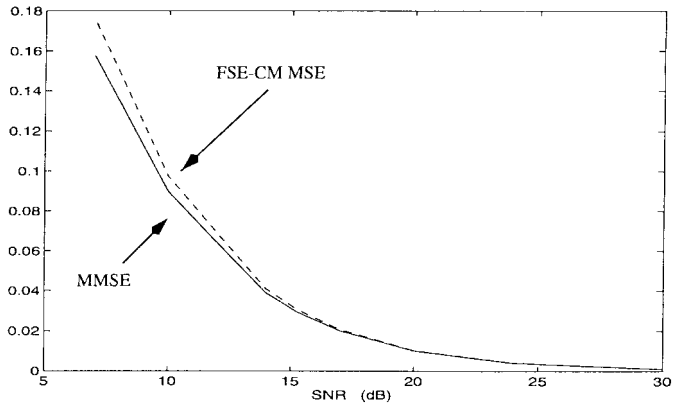



Fig. 7. Under ZF conditions (same simulation conditions as in Fig. 2): Comparison between MMSE and FSE-CMA MSE.

CM extrema. Indeed, the MMSE is, by definition, smaller than the FSE-CM MSE; however, they do not realize *a priori* the same tradeoff between ZF and noise enhancement. Next, we compare in more detail the FSE-CM and MMSE expressions.

### C. MSE Under ZF Conditions

From Proposal (2), we can easily deduce that the distance to ZF is $\|h_\gamma - h_\nu\|^2 = \gamma^2\|\bar{h}_\nu\|^2 + o(\gamma^2)$. In fact, because of (18), most MSE results from noise enhancement and can be expressed as

$$\gamma\|\vec{e}_\nu\|^2 = \gamma h_\nu^\top(\mathbf{C}^\top\mathbf{C}^*)^{-1}h_\nu + o(\gamma) \tag{19}$$

which is proportional to SNR. To have a more accurate expression of MSE, one needs to have a first-order expression of $\vec{e}$ in terms of $\gamma$. The first-order FSE-CM MSE is exactly the same as the first-order MMSE deduced in [8] by minimization of the MMSE criterion (4). Therefore, even if the expressions of the channel-equalizer global impulse responses differ between the MMSE and FSE-CM criterions minima, the corresponding performances are similar. As SNR decreases, we should see the FSE-CM MSE be the greater of the two, and we should also see that the difference grows. A remaining question is how to choose the equalizer so that the value of $\nu$ is achieved where the above Rayleigh ratio is minimized. In particular, if one wants to approach the lower MMSE bound when actually executing the algorithm in (6), a major open problem is the choice of the equalizer initialization so that the equalizer converges toward the global channel/equalizer with optimal delay $\nu$.

*1) Simulations:* We use model a) defined previously. We check the validity of the MSE analytical expression in terms of $\gamma$. Moreover, we want to see how close the vector $h_\gamma$ is to ZF. To make this latter measurement, the usual ISI measure is also used. The simulations on Fig. 6 show that $h_\gamma$ is close to a canonical vector ($\mu_{\text{ISI}} \ll 1$) even for a relatively small SNR. The FSE-CM MSE given by the analytical expression (19) is shown by $(-)$. It is compared with the FSE-CM criterion MSE $(--)$ estimated by averaging $(y(n) - s(n-\nu))^2$ at the criterion minimum setting $h_\gamma$ obtained when simulating the FSE-CM algorithm (6).

In Fig. 7, we compare the experimental MMSE $(-)$ and the experimental FSE-CM MSE $(--)$ on model a). The difference appears only for small SNR values (less than 10 dB), emphasizing the very good performance of the FSE-CM criterion in the noisy case.

### D. MSE Under Lack of Disparity

From Proposal 3 and under the same assumption on the equalizer length, i.e., $\mathbf{C}_0^\top d_\nu \approx h_\nu$, the distance to ZF is similar to the one in the case of achievable ZF.

$$\|h_\gamma - h_\nu\|^2 \approx \gamma^2\|\bar{h}_\nu\|^2 + o(\gamma^2)$$

the full expression of which is quite complex and of no particular interest. The error corresponding to noise enhancement is given by

$$\gamma\|\vec{e}_\nu\|^2 = \gamma h_\nu^\top \mathbf{C}_0^{-L*\top}(\underline{\mathbf{C}}^{*\top}\underline{\mathbf{C}})^{-1}\mathbf{C}_0^{-L}h_\nu + o(\gamma) \tag{20}$$

which is similar to expression (19) obtained in the case of ZF conditions when $h_\nu$ is replaced by $\mathbf{C}_0^{-L}h_\nu$. Equation (20) is the first-order approximation of the FSE-CMA MSE in the case of "long enough" equalizers. Comparing the MMSE lower bound, we see that (20) appears in the expression of MMSE when the zero-order term $\|(I - \mathbf{\Pi}_0)h_\nu\|^2$ is neglected. Next, we discuss the validity of (20) to estimate the MSE.

*1) Simulations:* We simulate equalizer lengths $N = 8$ (Fig. 8) and $N = 2$ (Fig. 9) with the multichannel model b). For $N = 8$, the equalizer is long enough to have a noise-free solution very close to a canonical vector ($\|(I - \mathbf{\Pi}_0)h_\nu\| \approx 0$) (Fig. 8). Consequently, (20) is approximatively valid for "large enough" equalizer length. For $N = 2$, the equalizer length is not large enough (four taps); therefore, in Fig. 9, an irreducible error appears in the experimental MSE $(-)$ even for very large SNR. We represent with $(.-)$ the nonavoidable error calculated as $\|(I - \mathbf{\Pi}_0)h_\nu\|^2$. If this measure is added to the analytical MSE curve in $(--)$, which corresponds to (20), we obtain a result that becomes very close to experimental MSE $(-)$.

This suggests that the FSE-CM MSE can be approximated, in terms of $\gamma$, by

$$\|(I - \mathbf{\Pi}_0)h_\nu\|^2 + \gamma h_\nu^\top \mathbf{C}_0^{-L*\top}(\underline{\mathbf{C}}^\top\underline{\mathbf{C}})^{-1}\mathbf{C}_0^{-L}h_\nu + o(\gamma) \tag{21}$$

which is exactly the first order in the $\gamma$ expression of the MMSE.

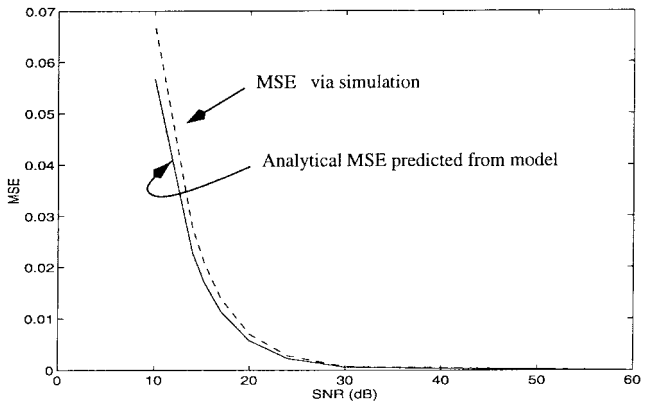



Fig. 8. Lack of disparity (simulation conditions as in Fig. 4): Impact of SNR on FSE-CMA MSE. Comparison between the simulated value and the analytical model.

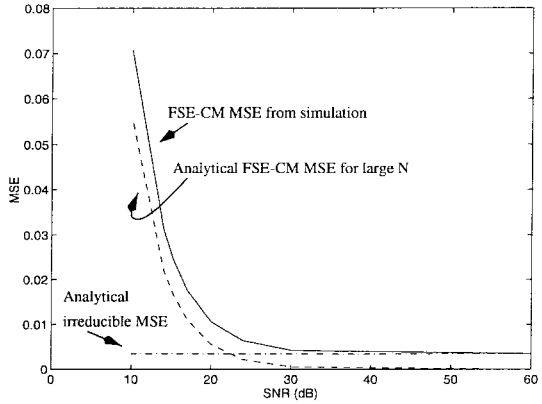



Fig. 9. Lack of disparity (same simulation conditions as in Fig. 4, except equalizer length $N = 2$): Impact of SNR on FSE-CMA MSE. Comparison between the simulated value, the analytical model for large $N$, and the analytical irreducible value.

## V. Conclusion

In this paper, we have analyzed the robustness of the FSE-CM criterion in the presence of additive receiver noise, whether or not the channel satisfies the ZF conditions. Expressing the additive noise's effect on the cost function, we have shown that noise introduces a smoothing effect on the FSE-CM cost function, which sets as a constraint on the output noise enhancement. By viewing the FSE-CM minima as a perturbation of the noise-free case, we produce an analytical model for the algorithm's asymptotical behavior in the limit of large SNR values. Moreover, the comparison (analytically and by means of simulations) between the resulting FSE-CM input-output MSE and the lower MMSE bound shows the following:

- Under ZF conditions, the FSE-CM criterion is doing almost as well as the MMSE criterion. In both cases, the main loss of performance is due to noise enhancement and not residual ISI.
- When there is loss of disparity, equalizers of sufficient length almost achieve ZF. In this case, the main MSE contribution is again due to noise enhancement.

However, when ZF is not achievable, the main contribution to FSE-CMA MSE is due to the irreducible MMSE. In both cases, the FSE-CM MSE is greater than but close to the MMSE or the equalizability lower bound. These results are corroborate by the study [24], where the FSE-CM minima are viewed as a perturbation of the MMSE ones.

With these results accomplished, the next step is the study of the FSE-CM Algorithm as a stochastic algorithm, including the excess MSE due to the stochastic jitter around the mean solution. Another important effort is a study of the algorithm initialization since it will be crucial to the attainment of the global channel/equalizer with optimal performance, i.e., the one with optimal value of $\nu$, which minimizes the MMSE.

## Appendix A
## Convolution Matrix Factorization

When the FIR subchannels transfer functions $c_1(z), \ldots, c_L(z)$ share $Z_0$ common zeros, the global impulse response of the channel equalizer system may be written as

$$h(z) = \vec{c}(z)^\top \vec{e}(z) = c_0(z)\, \underline{\vec{c}}(z)^\top \vec{e}(z) = c_0(z)\underline{e}(z)$$

where $\underline{e}(z) = \underline{\vec{c}}(z)^\top \vec{e}(z)$ is a FIR impulse response of length $N + Q - Z_0$. Then, one can associate with $d(z)$ the impulse response $\vec{d} = \underline{\mathbf{C}}^\top \vec{e}$. $\vec{e}$ is the $NL$-long impulse response associated with the equalizer $\vec{e}(z)$, and $\underline{\vec{c}}$ is a $(N + Q - Z_0)$-long impulse response. Thus, $\underline{\mathbf{C}}$ is a convolution matrix of dimension $NL \times (N + Q - Z_0)$. For the remaining part of the channel transfer function $c_0(z)$, one can associate the scalar transfer functions $h(z) = c_0(z)d(z)$ with the relation $h = \mathbf{C}_0^\top \vec{d}$, where $h$ is the $(N + Q)$-long global impulse response of the system. $\mathbf{C}_0$ is the $(N + Q - Z_0) \times (N + Q)$ channel convolution (Sylvester) matrix associated with $c_0(z)$. Note that the factorization of the multichannel convolution matrix is turned on the product of two Sylvester matrices.

## Appendix B
## First Minimization Step

The minimum value of $\Phi(\vec{e}) = K(h)\|\vec{e}\|^2 + \gamma\rho_g\|\vec{e}\|^4$, such as $h = \mathbf{C}^\top \vec{e}$, is given ($K(h) = 2(3\|h\|^2 - \rho)$ is viewed as a constant) as a constrained minimization problem that can be easily calculated by the Lagrange multiplier technique. Let $\lambda$ be a column vector with the same length as $h$. The first step in the Lagrange multiplier technique is to minimize the following real-valued expression versus $\vec{e}$.

$$\Phi(\vec{e}) + \lambda^\top(h - \mathbf{C}^\top \vec{e}) + (\lambda^\top(h - \mathbf{C}^\top \vec{e}))^*. \qquad (22)$$

For a given $h$, its derivation with respect to $\vec{e}$ gives $(K(h) + \gamma\rho_g\|\vec{e}\|^2)\vec{e}^* - \mathbf{C}^*\lambda^*$, which should be equal to zero (the minimum value), and results from solving $(K(h) + \gamma\rho_g\|\vec{e}\|^2)^2\|\vec{e}\|^2 = \lambda^\top \mathbf{C}^\top \mathbf{C}^*\lambda^*$, which is feasible but does not give much insight to the solutions.

As a first-order approximation of $J_\gamma(\vec{e})$ (in terms of $\gamma$), we will consider $\Phi(\vec{e}) = K(h)\|\vec{e}\|^2 + o(1)$. Thus, the minimization in (22) is achieved by $\vec{e} = (1/K)\mathbf{C}\lambda + o(1)$. Replacing $\vec{e}$ by this expression in (22) yields the following expression: $-(1/K)\lambda^\top \mathbf{C}^\top \mathbf{C}^*\lambda^* + \lambda^\top h + (\lambda^\top h)^*$ to be maximized versus $\lambda$. The desired $\lambda$ also satisfies $\mathbf{C}^\top \mathbf{C}^*\lambda^* = Kh$.

### A. Under ZF Conditions

Under the ZF condition, $\mathbf{C}$ is full column-rank; therefore, the matrix $(\mathbf{C}^\top\mathbf{C})$ is invertible. Thus, $\lambda = K\ (\mathbf{C}^\top\mathbf{C}^*)^{-1}\ h$, and $\vec{e} = (K/K)\ \mathbf{C}^*(\mathbf{C}^\top\mathbf{C}^*)^{-1}\ h + o(1)$. $\mathbf{C}^*(\mathbf{C}^\top\mathbf{C}^*)^{-1}\ h$ is the orthogonal projection of any $\vec{e}$ such that $h = \mathbf{C}^\top\vec{e}$ on the $(NL - (N+Q))$-dimensional subspace spanned by the columns of $\mathbf{C}^*$. It is also the result of the minimization procedure when considering the MMSE criterion (i.e., $\Phi(\vec{e}) = ||\vec{e}||^2$ [8]).

### B. Under Lack of Disparity

Here, $\mathbf{C} = \underline{\mathbf{C}}\mathbf{C}_0$; therefore, $\mathbf{C}^\top\mathbf{C}^*\lambda = Kh$ is equivalent to $\mathbf{C}_0\lambda = K\ (\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}(\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*\ h$. Therefore, $\vec{e} = (1/K)\mathbf{C}\lambda + o(1) = \underline{\mathbf{C}}^*(\underline{\mathbf{C}}^\top\underline{\mathbf{C}}^*)^{-1}(\mathbf{C}_0^*\mathbf{C}_0^\top)^{-1}\mathbf{C}_0^*\ h + o(1)$. The same remarks can be made as under the ZF conditions.

## ACKNOWLEDGMENT

The authors would like to thank O. Macchi and C. R. Johnson, Jr. for fruitful advice and discussions.

## REFERENCES

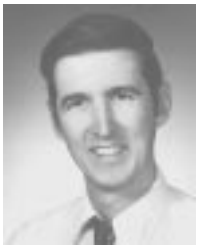
[1] K. Abed Meraim *et al.*, "Prediction error methods for time-domain blind identification od multichannel FIR filters," in *Proc. ICASSP'95*, 1995, pp. 1968–1971.
[2] R. R. Bitmead, S.-Y. Kung, B. D. O. Anderson, and T. Kailath, "Greatest common divisors via generalized Sylvester and Bezout matrices," *IEEE Trans. Automat. Contr.*, vol. 23, no. 6, pp. 1043–1047, 1978.
[3] Z. Ding, C. R. Johnson Jr., and R. A. Kennedy, "On the (non)existence of undesirable equilibria of Godard blind equalizers," *IEEE Trans. Signal Processing*, vol. 40, pp. 1940–1944, 1993.
[4] Z. Ding and Y. Li, "On channel identification based on second-order cyclic spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 1260–1264, May 1994.
[5] ———, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 43, no. 9, pp. 2120–2128, Sept. 1995.
[6] I. Fijalkow, "Multichannel equalization lower bound: A function of channel noise and disparity," in *Proc. IEEE SSAP-96*, 1996.
[7] I. Fijalkow, F. Lopez de Victoria, and C. R. Johnson, Jr., "Adaptive fractionally spaced blind equalization," in *Proc. IEEE DSP-94*, 1994.
[8] I. Fijalkow, A. Touzni, and C. R. Johnson, Jr., "Spatio-temporal equalizability under channel noise and loss of disparity," in *Proc. GRETSI*, 1995, vol. 1, pp. 293–296.
[9] I. Fijalkow, J. R. Treichler, and C. R. Johnson, Jr., "Fractionally spaced blind equalization: Loss of channel disparity," in *Proc. ICASSP'95*, 1995, vol. 4, pp. 1988–1991.
[10] D. Gesbert, P. Duhamel, and S. Mayrargue, "Subspace-based adaptive algorithms for the blind equalization of multichannel FIR filters," in *Proc. EUSIPCO'94*, 1994, vol. 2, pp. 712–715.
[11] G. B. Giannakis and S. D. Halford, "Blind fractionally-spaced equalization of noise FIR channels: Adaptive and optimal solutions," in *Proc. ICASSP'95*, 1995, vol. 4.
[12] D. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, 1980.
[13] C. R. Johnson, Jr., "Admissibility in blind adaptive channel equalization," *IEEE Contr. Syst. Technol.*, pp. 3–15, Jan. 1991.
[14] C. R. Johnson Jr., and B. D. O. Anderson, "Godard blind equalizer error surface characteristics: White, zero-mean, binary source case," *Int. J. Adaptive Contr. Signal Processing*, 1995.
[15] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
[16] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind identification of multichannel FIR systems," in *Proc. Int. Conf. Acoust. Speech Signal Processing*, 1994, vol. 4, pp. 581–584.
[17] S. Mayrargue, "A blind spatio-temporal equalizer for radio-mobile channel using the constant modulus algorithm (CMA)," in *Proc. ICASSP'93*, 1993, pp. 344–347.
[18] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, Jan. 1995.
[19] D. T. M. Slock and C. B. Papadias, "Further results on blind identification and equalization of multiple FIR channels," in *Proc. ICASSP'95*, 1995, vol. 4, pp. 1964–1967.
[20] L. Tong, G. Xu, and T. Kailath, "Blind identification and based on second-order statistics: A time domain approach," *IEEE Trans. Inform Theory*, vol. 40, no. 2, pp. 340–349, 1994.
[21] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind channel identification based on second order statistics: A frequency-domain approach," *IEEE Trans. Inform Theory*, vol. 41, Jan. 1995.
[22] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 2, pp. 459–472, 1983.
[23] J. R. Treichler, I. Fijalkow, and C. R. Johnson, Jr., "Fractionally spaced equalizers: How long should they be?," in *IEEE Signal Processing Mag.*, May 1996.
[24] H. Zeng and L. Tong, "Relationships between the CMA and MMSE receivers," in *Proc. CISS-96*, Mar. 1996.

**Inbar Fijalkow** was born in Haifa, Israel, on October 17, 1966. She received the Engineering and Ph.D. degrees from Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, June 1990 and October 1993, respectively.

She spent the 1993–1994 academic year at Cornell University, Ithaca, NY, as an Research Associate in the School of Electrical Engineering. Since September 1994, she is Maître de Conférences at Ecole Nationale Superieure de l'Electronique et de ses Applications (ENSEA), Cergy, France. She is also part of the Groupe De Recherche en Image, Signal et Informations (GDR ISIS) coordinated research project on blind equalization managed by the French Research Council (CNRS). Her current research interests are in the area of adaptive signal processing, particularly as applied to digital communications problems.

**Azzèdine Touzni** was born on July 12, 1969, in Paris, France. He received the Engineering degree from the Institut Galilé, Paris, France, in June 1994 and the Master degree from the Université de Cergy-Pontoise, France. He is currently working toward the Ph.D. degree from the Ecole Nationale Superieure de l'Electronique et de ses Applications (ENSEA), Cergy, France, in signal processing.

His research interests are focused on adaptive signal processing.



**John R. Treichler** (S'69–M'78–SM'89–F'92) was born in Velasco, TX, on September 22, 1947. He received the B.A. and Masters of Electrical Engineering degrees from Rice University, Houston, TX, in May 1970 and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in June 1977.

From 1970 to 1974, he was in the U.S. Navy and then spent the next three years at Stanford University. From 1977 to 1983, he was employed by ARGOSystems, Inc., Sunnyvale, CA. During this interval, he also served as a Lecturer at Stanford and then spent the 1983 to 1984 academic year at Cornell University, Ithaca, NY, as an Associate Professor in the School of Electrical Engineering. In 1984, he co-founded Applied Signal Technology, Inc., Sunnyvale, CA, where he still serves as the company's Senior Scientist. His current research interests are in the areas of digital and adaptive signal processing, particularly as applied to communications problems.