On-Board Regeneration of Uplink Signals Using a Blind Multichannel Estimator

MASSIMILIANO (MAX) MARTONE, Member, IEEE
Watkins-Johnson Co.

We propose a restoration method to compensate distortions caused by cross-polarization interference (CPI) and multipath propagation when multichannel transmission is employed in the uplink of an on-board-processing (OBP) satellite. The proposed baseband signal processing architecture is numerically robust and efficient as vector operations are avoided by efficient orthogonal transformations. Since the algorithm is based on higher-than-second-order statistics, the method is very effective in severe distortion conditions where a traditional, perfectly trained Kalman multichannel filter obtains poor performance.

Manuscript received December 15, 1994; revised October 1, 1996.

IEEE Log No. T-AES/34/1/00177.

This work was in part presented at GLOBECOM 95.

Author's address: Watkins-Johnson Co., Telecommunications Group, 700 Quince Orchard Rd., Gaithersburg, MD 20878-1794.

I. INTRODUCTION

On-board processing (OBP) is a technique for commercial satellite communications with various highly desirable features. At different levels of complexity and cost, digital technology allows the demodulation, decoding, switching, modulation and encoding to be performed on-board the satellite. Though a formidable technologic challenge, this will provide high efficiency interconnection for multiple access schemes and independent choice for signal modulation and encoding functions in the uplink and downlink. Specifically it is possible to implement real-time adaptive filtering to improve signal quality and compensate the deleterious propagation on the uplink path. Such enhancement will result in a more efficient communication system in terms of RF power utilization and in a considerable increase of the capacity of the multiple access scheme employed. One well known and effective method to reuse the available bandwidth in a satellite communication system is the transmission of modulated signals by means of orthogonally polarized waves. Channel distortions due to atmospheric propagation (for example rainfall) and antenna imperfections cause the two waves to be no longer orthogonally polarized so that the received signals are deteriorated by cross-polarization interference (CPI). CPI in a dually polarized satellite channel has characteristics that are randomly varying in time and are not known *a priori*: for example variations having a bandwidth of about 10 Hz are experienced in rainfall. Usual countermeausures [1] are based on the estimation of the scattering matrix (of the propagation channel) and the introduction of a differential phase shift and attenuation by means of RF networks to restore the wave orthogonality. Two main ideal assumptions in this approach are hardly satisfied in practical systems. First, since the estimation of the scattering matrix is performed using two pilot tones, this implies that the pilot tones can be perfectly separated from the desired information bearing signals. Second, the pilot tones are assumed to experience exactly the same distorting effects as the signals of interest.

Baseband digital signal processing is a more effective and reliable approach to mitigate CPI. This method was proposed in [2] where the feasibility and satisfactory results of the baseband approach were shown. The limitation of the algorithms of [2] was in the identification capability of the mean square error processor (or the channel estimator in the case of the maximum likelihood estimator). The minimum mean square error (MMSE) criterion is inherently based on second-order statistics (SOS). Satisfactory results can be obtained only when the equivalent vector-channel is (multi-)minimum-phase [3], and when the underlying random processes governing the time-varying characteristics of the system are assumed

to be Gaussian. However, most channels in the real world are not minimum-phase and their time-varying statistics are considerably different from the Gaussian assumption. In addition the MMSE approach requires a training signal (known at the receiver) which is not available.

Approaching the issue as a blind detection problem means that no knowledge of channels or transmitted signals is available. Blind deconvolution of nonminimum phase channels using higher order statistics (HOS) has stimulated considerable research effort over the past two decades; however, the generalization to the multichannel case is rarely practical, mainly because of the increased computational cost. In a recent paper Shalvi and Weinstein [4] proposed an algorithm to solve the blind deconvolution scalar problem using an iterative procedure. Here we present the generalization of this algorithm [4] to the multichannel case, and introduce a new efficient adaptive implementation based on a multichannel lattice filter. The algorithm also mitigates the effect of multipath propagation, which can be an important degradation factor in mobile satellite communications.

In Section II we give the necessary equations for the general model useful in the problem formulation, in Section III and Section IV we derive the block processing algorithm and in Section V we describe the adaptive implementation. The results of extensive computer simulations are presented in Section VI.

II. SYSTEM MODEL

Consider a dually polarized satellite digital radio communications system. Each channel supports two independent modulated signals (QPSK or BPSK, however we observe that the method is valid also for higher order M-QAM modulation schemes). The generated signal can be represented as

$$\begin{bmatrix} s^{(v)}(t) \\ s^{(h)}(t) \end{bmatrix} = e^{j\omega_0 t} \sum_n \begin{bmatrix} a_n^{(v)} \\ a_n^{(h)} \end{bmatrix} \mathbf{p}_{tx}(t - nT)$$

which is a modulation by two linearly polarized carrier waves in quadrature where $(\cdot)^{(v)}$ stands for the vertically polarized channel and $(\cdot)^{(h)}$ stands for the horizontally polarized channel, $a_n^{(v)}, a_n^{(h)}$ are the complex symbols, ω_0 is the carrier frequency, T is the signaling interval and

$$\mathbf{p}_{tx}(t) = \begin{bmatrix} p_v(t) & p_{h,v}(t) \\ p_{v,h}(t) & p_h(t) \end{bmatrix}$$

is the transmitter (matrix) filter whose Fourier transform is $\mathcal{P}_{tx}(\omega)$. The diagonal entries of $\mathcal{P}_{tx}(\omega)$ are transmitter impulse responses acting as signal shapers, while the off-diagonal impulse responses act as cross-channel signal shapers and power distributors.

We work with the baseband complex representation of the signals and adopt the following notation: vectors and matrices are bold, $\mathbf{M}^T$, $\mathbf{v}^T$ designating transposition for matrix $\mathbf{M}$ and vector $\mathbf{v}$, respectively. Complex conjugation for scalars, matrices, vectors is indicated as u^*, $\mathbf{M}^*$, $\mathbf{v}^*$, respectively, while notations $[\mathbf{M}]_{l,m}$ and $[\mathbf{v}]_k$ stand for the l,m element of matrix $\mathbf{M}$ and the kth element of vector $\mathbf{v}$, respectively. Finally, $\mathbf{I}_m$ is the $m \times m$ identity matrix and the symbol " * " means convolution.

The channel is characterized in the frequency domain by the complex matrix

$$\mathcal{C}(\omega) = \begin{bmatrix} \mathcal{F}_v(\omega) & \mathcal{F}_{h,v}(\omega) \\ \mathcal{F}_{v,h}(\omega) & \mathcal{F}_h(\omega) \end{bmatrix}$$

where $\mathcal{F}_v(\omega)$, $\mathcal{F}_h(\omega)$, are Fourier transforms of

$$f_v(t) = \sum_{m=1}^{N_v} \rho_{v,m} e^{j\psi_{v,m}} \delta(t - \tau_{v,m})$$

$$f_h(t) = \sum_{m=1}^{N_h} \rho_{h,m} e^{j\psi_{h,m}} \delta(t - \tau_{h,m})$$

with $\tau_{v,m}$, $\rho_{v,m}$, $\psi_{v,m}$, N_v, $\tau_{h,m}$, $\rho_{h,m}$, $\psi_{h,m}$, N_h delay, amplitude, phase and number of rays of the multipath channel relative to the vertically and horizontally polarized channel, respectively [5]. $\delta(t)$ is the delta function. The characterization of $\mathcal{F}_{h,v}(\omega)$, $\mathcal{F}_{v,h}(\omega)$ is extremely difficult, but based on some previous studies [6, 7], we characterize them as follows

$$\mathcal{F}_{h,v}(\omega) = K_1 \mathcal{F}_v(\omega) + K_2 \mathcal{F}_h(\omega) + R_1 e^{j\omega D_1}$$

$$\mathcal{F}_{v,h}(\omega) = K_3 \mathcal{F}_v(\omega) + K_4 \mathcal{F}_h(\omega) + R_2 e^{j\omega D_2}$$

where K_1, K_2, K_3, K_4 are constants that incorporate the nonideal properties of the antennas at both ends of the channels. The exponential terms are nondispersive cross-polarization responses contributed by an independent ray.

We are assuming for now a time invariant channel. This condition holds in many applications of interest since the observation interval is often much shorter than the coherence time of the channel which characterizes the time-variant behavior of the propagation media. However, the adaptive scheme described in Section V is designed for time-varying channels. We assume that noise is additive white Gaussian on the vertical and horizontal channel and is characterized by two independent complex processes $\eta^{(v)}(t)$, $\eta^{(h)}(t)$ such that $E\{\mathbf{n}(t)\mathbf{n}(s)^{T^*}\} = N_0 \delta(t-s)\mathbf{I}_2$, where $\mathbf{n}(t) = (\eta^{(v)}(t), \eta^{(h)}(t))^T$. In matrix notation we can write

$$\mathbf{r}(t) = \int_{-\infty}^{+\infty} \mathbf{c}(t - \tau)\mathbf{s}(\tau)\, d\tau + \mathbf{n}(t) \qquad (1)$$

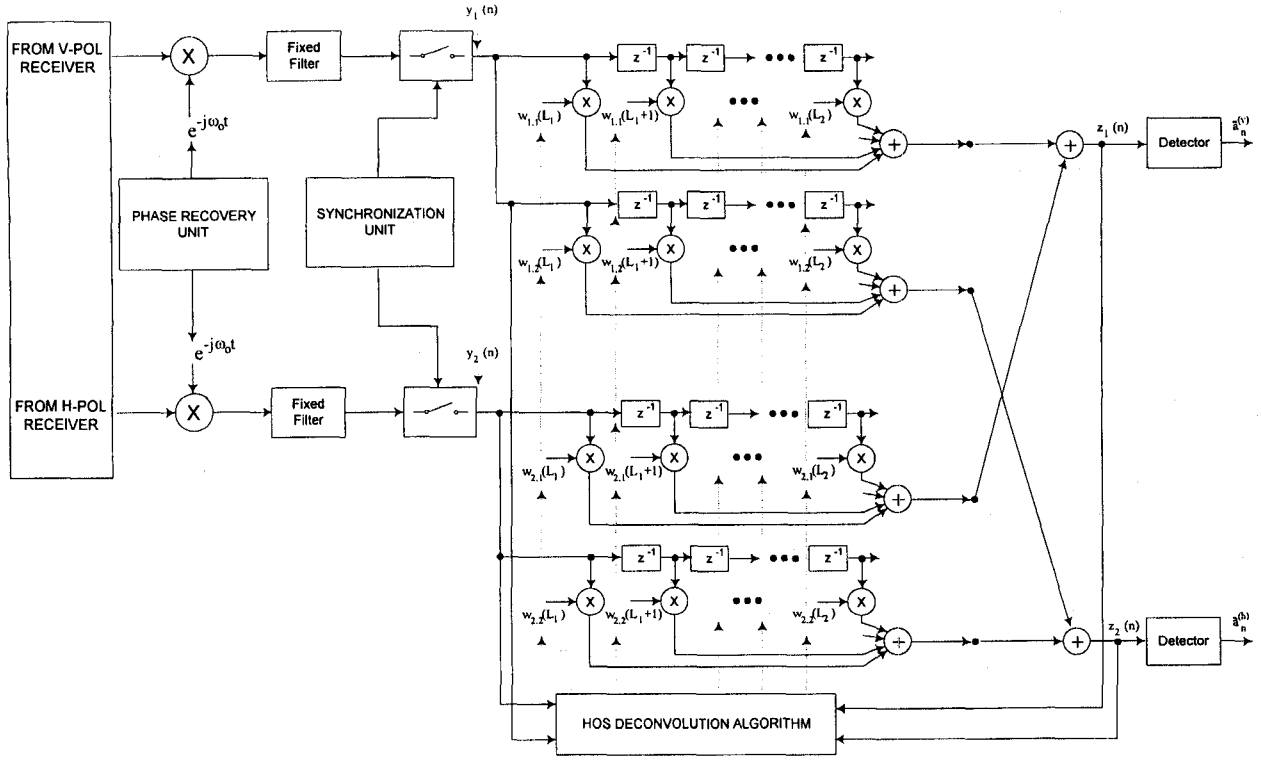


Fig. 1. Block diagram of signal processing section of receiver, ω_0 is carrier frequency. Multiplications and additions are complex operations.

having denoted $\mathbf{s}(t) = [s^{(v)}(t), s^{(h)}(t)]^T$, $\mathbf{r}(t) = [r^{(v)}(t), r^{(h)}(t)]^T$, and

$$\mathbf{c}(t) = \begin{bmatrix} f_v(t) & f_{h,v}(t) \\ f_{v,h}(t) & f_v(t) \end{bmatrix}.$$

The average power transmitted on the vertically polarized channel is

$$P_v = \frac{\sigma_x^2}{T} \int_{-\infty}^{+\infty} (|p_v(t)|^2 + |p_{h,v}(t)|^2)\, dt$$

and on the horizontally polarized channel it is

$$P_h = \frac{\sigma_x^2}{T} \int_{-\infty}^{+\infty} (|p_h(t)|^2 + |p_{v,h}(t)|^2)\, dt.$$

In the following we assume that the transmit filters are square root raised cosine filters, that $\mathbf{p}_{tx}(t)$ is diagonal ($p_{v,h}(t) = p_{h,v}(t) = 0$). The receiver matrix filter $\mathbf{p}_{rx}(t)$ whose Fourier transform is $\mathcal{P}_{rx}(\omega)$ is also diagonal and constituted by filters perfectly matched to the transmitter filters. A Nyquist roll-off factor of 35 percent is assumed.

If we define

$$\mathbf{H}(m) = \begin{bmatrix} h_{1,1}(m) & h_{1,2}(m) \\ h_{2,1}(m) & h_{2,2}(m) \end{bmatrix} = \mathbf{p}_{rx}(t) * \mathbf{c}(t) * \mathbf{p}_{tx}(t)|_{t=mT}$$

we can compactly express the T-sampled discrete time system as

$$\mathbf{y}(n) = \sum_k \mathbf{H}(k)\mathbf{x}(n-k) + \mathbf{n}(n) \qquad (2)$$

where now $\mathbf{x}(n) = [x_1(n), x_2(n)]^T = [a_n^{(v)}, a_n^{(h)}]^T$, $\mathbf{y}(n) = [y_1(n), y_2(n)]^T = [r^{(v)}(nT), r^{(h)}(nT)]^T$, $\mathbf{n}(n) = [n_1(n), n_2(n)]^T = [\eta^{(v)}(nT), \eta^{(h)}(nT)]^T$. The two independent input data streams have covariance matrix $E\{\mathbf{x}(n)\mathbf{x}(n)^{T^*}\} = \mathrm{diag}(\sigma_x^2, \sigma_x^2)$.

III. DISTORTIONLESS RECEPTION

In the z-domain the vector impulse response $\mathbf{H}(m)$ in (2) can be expressed as

$$\tilde{\mathcal{H}}(z) = \sum_{k=-\infty}^{\infty} \mathbf{H}(k) z^{-k} = \begin{bmatrix} \mathcal{H}_{1,1}(z) & \mathcal{H}_{1,2}(z) \\ \mathcal{H}_{2,1}(z) & \mathcal{H}_{2,2}(z) \end{bmatrix}.$$

We concentrate on the distorting effects of $\tilde{\mathcal{H}}(z)$ on the signals and discard the effects of the additive Gaussian noise $\mathbf{n}(n)$ for the moment. To recover the input signals a linear 2-input 2-output filter

$$\tilde{\mathcal{W}}(z) = \sum_{k=L_1}^{L_2} \mathbf{W}(k) z^{-k} = \begin{bmatrix} \mathcal{W}_{1,1}(z) & \mathcal{W}_{1,2}(z) \\ \mathcal{W}_{2,1}(z) & \mathcal{W}_{2,2}(z) \end{bmatrix}$$

with length $L = L_2 - L_1 + 1$ is applied to the outputs of the 2-channel model $\mathcal{H}(z)$ as depicted in the block diagram of Fig. 1 where we used complex representation of signals. We assume ideally perfect phase and timing recovery. The algorithm we propose uses the observations $y_1(n)$, $y_2(n)$ and the outputs of the filter $z_1(n)$ $z_2(n)$ to derive at each time step the new set of weights. While the statistics of $y_1(n)$, $y_2(n)$

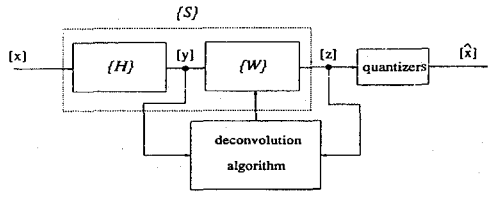


Fig. 2. Mathematical model for deconvolution problem.

are time-invariant (or at least slowly time-variant) because $\tilde{\mathcal{H}}(z)$ is invariant, the statistics of $z_1(n)$ $z_2(n)$ are time-varying because the filter $\tilde{\mathcal{W}}(z)$ is updated at every step until it converges to the inverse of $\tilde{\mathcal{H}}(z)$ (an approximation of its inverse, in practice). From a mathematical point of view the deconvolution problem we have to solve is shown in Fig. 2.

The recent tremendous progress in hardware technology allows the implementation of all the functions in the block diagram of Fig. 1 to be performed by digital circuits. Particularly the baseband downconversion can be implemented using digital downconverters. The advantage of the digital implementation stems from the improvement of accuracy and stability, increased flexibility, and considerable cost reduction of the overall communication equipment on board the satellite.

The main objective for $\tilde{\mathcal{W}}(z)$ is *distortionless reception*, that is formally

$$\tilde{\mathcal{W}}(z)\tilde{\mathcal{H}}(z) = \mathbf{I}_2 \tag{3}$$

where $\mathbf{I}_2$ is a 2×2 identity matrix. The system $\tilde{\mathcal{W}}(z)$ is required to be bounded-input bounded-output (BIBO) stable. The solution (3) is achievable only in ideal cases. Since the input signal constellations are symmetric, the statistics of the input signals $x_i(n)$ reflect the same symmetry. Moreover, *signal reconstruction is possible only up to a constant delay, due to the stationarity of the input process*. The recovered signals will be subject to a phase ambiguity, a delay, and a permutation ambiguity [9, 10]. The best possible result for practical *distortionless reception* by means of a linear filter is

$$\tilde{\mathcal{W}}(z)\tilde{\mathcal{H}}(z) = \mathbf{P}\mathcal{D}(z) \tag{4}$$

where $\mathbf{P}$ is a permutation matrix and

$$\mathcal{D}(z) = \operatorname{diag}\{e^{j\phi_1}z^{-n_1}, e^{j\phi_2}z^{-n_2}\}$$

$$\phi_i \in [-\pi, \pi], \qquad n_i \text{ integer for } \quad i = 1,2.$$

In the time domain the linear filter can be written as

$$z_i(n) = \sum_{j=1}^{2}\sum_{m} w_{i,j}(m)y_j(n-m), \qquad i = 1,2 \tag{5}$$

where $w_{i,j}(m)$ is the filter corresponding to the polynomial $\mathcal{W}_{i,j}(z)$. The overall system is characterized by the input/output relation

$$z_i(n) = \sum_{j=1}^{2}\sum_{m} s_{i,j}(m)x_j(n-m) \qquad i = 1,2.$$

The impulse response of the two cascaded filters is given by the multivariate convolution

$$s_{i,j_1}(m_2) = \sum_{j=1}^{2}\sum_{m} w_{i,j}(m)h_{j,j_1}(m_2 - m)$$

$$= \left[\sum_{j=1}^{2}\mathbf{H}_{j,j_1}\mathbf{w}_{i,j}\right]_{m_2} \qquad i, j_1 = 1,2$$

or in a vector form

$$\mathbf{s}_{i,j_1} = \sum_{j=1}^{2}\mathbf{H}_{j,j_1}\mathbf{w}_{i,j} \qquad i, j_1 = 1,2$$

and

$$\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i \tag{6}$$

where $\tilde{\mathbf{H}}$, $\tilde{\mathbf{s}}_i$, $\tilde{\mathbf{w}}_i$ are defined as

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_{1,1} & \mathbf{H}_{2,1} \\ \mathbf{H}_{1,2} & \mathbf{H}_{2,2} \end{bmatrix}$$

$$[\mathbf{H}_{i,j}]_{m,n} = h_{i,j}(m-n) \qquad -\infty \le m \le +\infty, \quad L_1 \le n \le L_2$$

$$\tilde{\mathbf{s}}_i = (\mathbf{s}_{i,1}^T, \mathbf{s}_{i,2}^T)^T \tag{7}$$

$$\mathbf{s}_{i,j} = (\ldots, s_{i,j}(-1), s_{i,j}(0), s_{i,j}(1), \ldots)^T$$

$$\tilde{\mathbf{w}}_i = (\mathbf{w}_{i,1}{}^T, \mathbf{w}_{i,2}{}^T)^T$$

$$\mathbf{w}_{i,j} = (w_{i,j}(L_1), w_{i,j}(L_1+1), \ldots, w_{i,j}(L_2))^T.$$

The desired solution $\tilde{\mathbf{s}}_i$ that completely restores the information signal of the ith channel up to the delay n_i, is defined as

$$\tilde{\delta}_i = (\delta_{i,1}^T, \delta_{i,2}^T)^T, \qquad \delta_{i,j} = \begin{cases} (\ldots, 0, 0, 0, \ldots) & i \neq j \\ \hat{\delta}_i & i = j \end{cases}.$$

The generic mth element of the vector $\hat{\delta}_i$ is $\delta(m - n_i)$, if we neglect the phase shift and we force the solution to not permute the inputs ($\mathbf{P} = \mathbf{I}_2$). This generalizes the one-dimensional case [4]. It is possible to solve this deconvolution problem by solving the minimization problem

$$\min_{\tilde{\mathbf{w}}_i} \|\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i - \tilde{\delta}_i\|^2. \tag{8}$$

We now make important assumptions on the discrete-time model described and detail some requirements on the statistical properties of the input symbols distribution.

The high-order statistical properties of a process are commonly described in the time domain by cumulants. Cumulants of interest here are fourth-order

cumulants of complex zero mean stationary processes $x(i)$ defined as [8]

$$\begin{aligned}\text{cum}[x(i_1),x(i_2),x^*(i_3),x^*(i_4)] &= E\{x(i_1)x(i_2)x(i_3)^*x(i_4)^*\} \\ &\quad - E\{x(i_1)x(i_2)\}E\{x(i_3)^*x(i_4)^*\} \\ &\quad - E\{x(i_1)x(i_3)^*\}E\{x(i_2)x(i_4)^*\} \\ &\quad - E\{x(i_1)x(i_4)^*\}E\{x(i_2)x(i_3)^*\}.\end{aligned}$$

The properties of cumulants that we exploit are as follows.

LIN: $\text{cum}[\sum_i f(i)x(i),\ldots] = \sum_i f(i)\text{cum}[x(i),\ldots,]$.
STATIND: if the samples of a process can be divided into two (or more) statistically independent subsets, then their joint cumulants are zero.

It is also well known that if the process samples are jointly Gaussian, then their joint kth-order cumulant is zero for $k > 2$.

The fundamental assumptions necessary to develop the algorithm are the following.

AS1: $\tilde{\mathcal{H}}(z)$ represents a stable system.
AS2: The inverse of the system $\tilde{\mathcal{H}}(z)$ exists and may be noncausal.
AS3: The complex sequence $\{x_i(n)\}$ is constituted by random variables identically non-Gaussian distributed and statistically independent, and the cumulants of $\{x_i(n)\}$ satisfy
$E\{x_i(n)\} = E\{x_i^3(n)\} = 0$.
$\text{cum}[x_i(n),x_j^*(n)] = \sigma_x^2 > 0$, only for $i = j$.
$E\{x_i(n)x_j(n)\} = E\{x_i^*(n)x_j^*(n)\} = 0$, for any i,j.
$\text{cum}[x_{j_1}(n),x_{j_2}(n),x_{j_3}^*(n),x_{j_4}^*(n)] = \text{cum}_4(x) \neq 0$, only for $j_1 = j_2 = j_3 = j_4$.

The solution of (8) is

$$\tilde{\mathbf{w}}_i = (\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\delta}_i \tag{9}$$

and $\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\delta}_i$.

IV. DERIVATION OF THE ALGORITHM

The following two-step iterative procedure defines a class of algorithms for different values of p and q [4]

$$s_{i,j}(k)^{[1]} = (s_{i,j}(k))^p(s_{i,j}(k)^*)^q \tag{10}$$

$$s_{i,j}(k)^{[2]} = s_{i,j}(k)^{[1]}\frac{1}{\sqrt{\sum_{j=1}^2\sum_l |s_{i,j}(l)^{[1]}|^2}} \tag{11}$$

where $(\cdot)^{[1]}$, $(\cdot)^{[2]}$ stand for the result of the first and second step, respectively. This iterative method converges at a "*super exponential*" rate to the desired solution for $p + q \geq 2$ [4]. Here we choose $p = 2$, $q = 1$, which gives a solution in terms of fourth-order cumulants. Since obviously $\tilde{\mathbf{s}}_i$ is not available, (because $\tilde{\mathbf{H}}$ is not known) we derive a procedure in terms of $\tilde{\mathbf{w}}_i$. If we define $\tilde{\mathbf{g}}_i = (\mathbf{g}_{i,1}^T, \mathbf{g}_{i,2}^T)^T$, with $\mathbf{g}_{i,j} = (\ldots, g_{i,j}(-1), g_{i,j}(0), g_{i,j}(1), \ldots)^T$, as the vector impulse response obtained by $g_{i,j}(k) = s_{i,j}^2(k)s_{i,j}(k)^*$, we can state the least squares minimization problem

$$\min_{\tilde{\mathbf{w}}_i} \|\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i^{[1]} - \tilde{\mathbf{g}}_i\|^2 \tag{12}$$

whose solution is

$$\tilde{\mathbf{w}}_i^{[1]} = (\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{g}}_i. \tag{13}$$

To obtain normalization (11), the second step is

$$\tilde{\mathbf{w}}_i^{[2]} = \frac{\tilde{\mathbf{w}}_i^{[1]}}{\sqrt{\tilde{\mathbf{w}}_i^{[1]^{T^*}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})\tilde{\mathbf{w}}_i^{[1]}}}. \tag{14}$$

A discussion on the convergence of (13)–(14), related to the convergence of (10)–(11) is given in Appendix B. The procedure (13)–(14) can be expressed in terms of the cumulants of the outputs of the sensors. From $y_i(k-n) = \sum_{j=1}^2\sum_m h_{i,j}(m-n)x_j(k-m)$, we exploit the above assumptions and the properties of the cumulants of linear stationary processes (see property LIN) so that we can write

$$\begin{aligned}&\text{cum}[y_{i_1}(k-n), y_{i_2}{}^*(k-m)] \\ &= \text{cum}\left[\sum_{j_1=1}^2\sum_{m_1} h_{i_1,j_1}(m_1-n)x_{j_1}(k-m_1), \right. \\ &\qquad \left. \sum_{j_2=1}^2\sum_{m_2} h_{i_2,j_2}^*(m_2-m)x_{j_2}^*(k-m_1)\right] \\ &= \sum_{j=1}^2\sum_k h_{i_1,j}(k-n)h_{i_2,j}^*(k-m)\sigma_x^2 \\ &= \sigma_x^2\sum_{j=1}^2(\mathbf{H}_{i_2,j}^{T^*}\mathbf{H}_{i_1,j})_{m,n}\end{aligned} \tag{15}$$

due to

$$\text{cum}[x_{j_1}(k-m_1), x_{j_2}^*(k-m_2)] = \begin{cases}\text{cum}[x_j(n),x_j^*(n)] = \sigma_x^2 & j_1 = j_2;\quad m_1 = m_2 \\ 0 & \text{otherwise}\end{cases}.$$

To derive the second key expression related to the solution (13) let us consider

$$\begin{aligned}&\text{cum}[z_i(k), z_i(k), z_i^*(k), y_{i_3}^*(k-n)] \\ &= \sum_{j=1}^2\sum_m h_{i_3,j}^*(m-n)\text{cum}[z_i(k), z_i(k), z_i^*(k), x_j^*(k-m)]\end{aligned}$$

and

$$\begin{aligned}&\text{cum}[z_i(k), z_i(k), z_i^*(k), x_j^*(k-m)]\\&= \sum_{j_1=1}^{2}\sum_{j_2=1}^{2}\sum_{j_3=1}^{2}\sum_{l_1}\sum_{m_1}\sum_{n_1} s_{i,j_1}(l_1)s_{i,j_2}(m_1)s_{i,j_3}^*(n_1)\\&\quad\times \text{cum}[x_{j_1}(k-l_1), x_{j_2}(k-m_1),\\&\qquad x_{j_3}^*(k-n_1), x_j^*(k-m)]\\&= \text{cum}_4(x)s_{i,j}^2(m)s_{i,j}^*(m)\end{aligned}$$

where the last equality follows from

$$\begin{aligned}&\text{cum}[x_{j_1}(n-m_1), x_{j_2}(n-m_2), x_{j_3}^*(n-m_3), x_j^*(n-k)]\\&= \begin{cases}\text{cum}[x_j(n), x_j(n), x_j^*(n), x_j^*(n)] = \text{cum}_4(x) & j_1 = j_2 = j_3 = j;\quad m_1 = m_2 = m_3 = k\\ 0 & \text{otherwise}\end{cases}.\end{aligned}$$

So we can write

$$\begin{aligned}&\text{cum}[z_i(k), z_i(k), z_i^*(k), y_{i_3}^*(k-n)]\\&= \sum_{j=1}^{2}\sum_{m} h_{i_3,j}^*(m-n)g_{i,j}(m)\text{cum}_4(x)\\&= \text{cum}_4(x)\sum_{j=1}^{2}(\mathbf{H}_{i_3,j}^{T^*}\mathbf{g}_{i,j})_n. \qquad (16)\end{aligned}$$

Expressions (15) and (16) can be substituted in the least squares solution (13) and the following iterative algorithm is obtained

$$\tilde{\mathbf{w}}_i^{[1]} = \tilde{\mathbf{R}}^{-1}\mathbf{D}_i \qquad (17)$$

$$\tilde{\mathbf{w}}_i^{[2]} = \frac{\tilde{\mathbf{w}}_i^{[1]}}{\sqrt{\tilde{\mathbf{w}}_i^{[1]T^*}\tilde{\mathbf{R}}\tilde{\mathbf{w}}_i^{[1]}}} \qquad (18)$$

where the matrix $\tilde{\mathbf{R}}$ with dimensions $2L \times 2L$, is given by

$$\tilde{\mathbf{R}} = \begin{bmatrix}\mathbf{R}_{1,1} & \mathbf{R}_{1,2}\\ \mathbf{R}_{2,1} & \mathbf{R}_{2,2}\end{bmatrix}$$

$$[\mathbf{R}_{i,j}]_{n,m} = \frac{\text{cum}[y_j(k-m), y_i^*(k-n)]}{\sigma_x^2}$$

and the $2L \times 1$ vector $\mathbf{D}_i$ of fourth-order cumulants is given by

$$\mathbf{D}_i = (\mathbf{d}_{i,1}^T, \mathbf{d}_{i,2}^T)^T,$$

$$[\mathbf{d}_{i,m}]_n = \frac{\text{cum}[z_i(k), z_i(k), z_i^*(k), y_m^*(k-n)]}{\text{cum}_4(x)}.$$

V. ADAPTIVE IMPLEMENTATION

In this section we derive an adaptive algorithm for on-line computation of the multichannel deconvolution filter parameters. A well-known result is that if the samples are well separated in time, and if the cumulants are absolutely summable, then theoretical cumulants are consistently estimated from a data record of N samples and ensemble averages can be approximated by empirical averages. We have for the cumulants of interest in the algorithm

$$\overline{\text{cum}}[y(k-n), y^*(k-m)] = \frac{1}{N}\sum_{k=1}^{N} y(k-n)y^*(k-m) \qquad (19)$$

and

$$\begin{aligned}&\overline{\text{cum}}[z(k), z(k), z^*(k), y^*(k-n)]\\&= \frac{1}{N}\sum_{k=1}^{N}|z(k)|^2 y^*(k-n)z(k)\\&\quad - 2\frac{1}{N}\sum_{k=1}^{N}|z(k)|^2\frac{1}{N}\sum_{k=1}^{N}z(k)y^*(k-n)\\&\quad - \frac{1}{N}\sum_{k=1}^{N}z(k)^2\frac{1}{N}\sum_{k=1}^{N}z^*(k)y^*(k-n) \qquad (20)\end{aligned}$$

where we have omitted indices for $y(k)$, $z(k)$ for simplicity, and we have indicated the estimated cumulant as $\overline{\text{cum}}[\cdot\cdot]$. At the end of the convergence process

$$\mathbf{D}_i - \tilde{\mathbf{R}}\tilde{\mathbf{w}}_i = \mathbf{0}$$

must be satisfied. We assume that the power constraint (14) is always satisfied using an AGC (automatic gain control). From the derivation it clearly follows that at each $t+1$ stage we wish to solve the problem

$$\min_{\tilde{\mathbf{w}}_i}\left\|\begin{pmatrix}\lambda\tilde{\mathbf{Y}}^{(n)}\\ \tilde{\mathbf{y}}(n+1)^T\end{pmatrix}\tilde{\mathbf{w}}_i - \begin{pmatrix}\lambda\mathbf{Z}_i^{(n)}\\ \tilde{z}_i(n+1)\end{pmatrix}\right\|^2 \qquad (21)$$

with

$$\begin{aligned}\mathbf{Z}_i^{(n)} &= (\tilde{z}_i(1), \tilde{z}_i(2), \ldots, \tilde{z}_i(n))^T\\(\tilde{\mathbf{Y}}^{(n)})^T &= (\tilde{\mathbf{y}}(1), \tilde{\mathbf{y}}(2), \ldots, \tilde{\mathbf{y}}(n))\\\tilde{\mathbf{y}}(n) &= (\breve{\mathbf{y}}_1(n)^T, \breve{\mathbf{y}}_2(n)^T)^T\\\breve{\mathbf{y}}_i(n) &= (y_i(n-L_1), y_i(n-L_1-1), \ldots, y_i(n-L_2))^T\end{aligned} \qquad (22)$$

and

$$\tilde{z}_i(n) = (|z_i(n)|^2 - 2\sigma_x^2)z_i(n). \qquad (23)$$

In fact, the *normal equations* define the desired minimizer as $\tilde{\mathbf{Y}}^{(n+1)^{T^*}}\tilde{\mathbf{Y}}^{(n+1)}\tilde{\mathbf{w}}_i = \tilde{\mathbf{Y}}^{(n+1)^{T^*}}\mathbf{Z}_i^{(n+1)}$

TABLE I
Dimension of Lattice Parameters

$\mathbf{e}_n^{(m)}$	2×1
$\mathbf{r}_n^{(m)}$	2×1
$e_n^{(m,i)^x}$	1×1
$\mathbf{C}_n^{e(m)}$	2×2
$\mathbf{C}_n^{r(m)}$	2×2
$\mathbf{\Delta}_n^{(m)}$	2×2
$(\gamma_n^{(m)})^2$	1×1
$\mathbf{\Delta}_n^{(m)^x}$	1×2

which is equivalent to $\mathbf{D}_i - \tilde{\mathbf{R}}\tilde{\mathbf{w}}_i = \mathbf{0}$, if we employ sample statistics estimators for the cumulants and the covariance matrix. Expression (21) can be justified by considering the estimation of fourth-order cumulants, based on sample average given by (20) and **AS3** (the third term on the right-hand side of (20) does not need to be estimated since $E\{z_i^2(n)\} = \sum_j \sum_k s_{i,j}(k)^2 E\{x_i(n)x_i(n)\} = 0$). Moreover, due to the power normalization (14),

$$E\{|z_i(n)|^2\} = E\{z_i(n)z_i^*(n)\} = \sum_j \sum_k |s_{i,j}(k)|^2 E\{x_i(n)x_i^*(n)\} = \sigma_x^2.$$

Expression (21) reveals a similarity with the standard recursive least squares (RLS) problem where the process $\tilde{z}_i(n)$ is the desired signal.

The Multichannel Lattice Deconvolution Filter: The derivation of the recursive expressions for the lattice predictor, not described here, can be made by an algebraic or a geometric approach (see for details [11–13]). The main modification with respect to the traditional lattice least squares filter is in the joint process estimation. The process $\tilde{z}_i(n)$, obtained from the output of the filter estimated at time step $n-1$, is at each time step the desired signal. The straightforward generalization of the single channel algorithm to the multichannel case is given in the following.

Superscripts denote order of the lattice stage ($m = L_1, L_1+1, \ldots, L_2$). Dimensions of the lattice parameters system are in Table I.

Backward/Forward Recursions

Forward Prediction Error:

$$\mathbf{e}_n^{(m)} = \mathbf{e}_n^{(m-1)} - (\mathbf{\Delta}_n^{(m)})^{T^*}(\mathbf{C}_n^{r(m)})^{-1}\mathbf{r}_{n-1}^{(m-1)}.$$

Backward Prediction Error:

$$\mathbf{r}_n^{(m)} = \mathbf{r}_{n-1}^{(m-1)} - (\mathbf{\Delta}_n^{(m)})^{T^*}(\mathbf{C}_n^{e(m)})^{-1}\mathbf{e}_n^{(m-1)}.$$

Forward Error Correlation:

$$\mathbf{C}_n^{e(m)} = \lambda^2 \mathbf{C}_{n-1}^{e(m)} + \frac{\mathbf{e}_n^{(m-1)}(\mathbf{e}_n^{(m-1)})^{T^*}}{(\gamma_{n-1}^{(m-1)})^2}.$$

Backward Error Correlation:

$$\mathbf{C}_n^{r(m)} = \lambda^2 \mathbf{C}_{n-1}^{r(m)} + \frac{\mathbf{r}_{n-1}^{(m-1)}(\mathbf{r}_{n-1}^{(m-1)})^{T^*}}{(\gamma_{n-1}^{(m-1)})^2}.$$

Forward/Backward Cross-Correlation:

$$\mathbf{\Delta}_n^{(m)} = \lambda^2 \mathbf{\Delta}_{n-1}^{(m)} + \frac{\mathbf{e}_n^{(m-1)}(\mathbf{r}_{n-1}^{(m-1)})^{T^*}}{(\gamma_{n-1}^{(m-1)})^2}.$$

Likelihood Variable:

$$(\gamma_n^{(m)})^2 = (\gamma_{n-1}^{(m-1)})^2 - (\mathbf{e}_n^{(m-1)})^{T^*}(\mathbf{C}_n^{e(m)})^{-1}\mathbf{e}_n^{(m-1)}.$$

Joint Process Recursions

Joint Estimation Error:

$$e_n^{(m,i)^x} = e_n^{(m-1,i)^x} - \mathbf{\Delta}(m)_n^x (\mathbf{C}_n^{r(m)})^{-1}\mathbf{r}_{n-1}^{(m-1)}.$$

Estimation Cross-Correlation:

$$\mathbf{\Delta}_n^{(m)^x} = \lambda^2 \mathbf{\Delta}_{n-1}^{(m)^x} + \frac{e_n^{(m-1,i)^x}(\mathbf{r}_{n-1}^{(m-1)})^{T^*}}{(\gamma_{n-1}^{(m-1)})^2}.$$

Initialization

$\mathbf{e}_n^{(L_1)} = (y_1(n-L_1), y_2(n-L_1))^T$, $\mathbf{r}_0^{(m)} = \mathbf{0}$, $e_n^{(L_1,i)^x} = \tilde{z}_i(n)$.
$\mathbf{r}_n^{(L_1)} = (y_1(n-L_1), y_2(n-L_1))^T$.
$(\gamma_0^{(m)})^2 = 1$, $(\gamma_n^{(L_1)})^2 = 1$.
$\mathbf{C}_0^{e(m)} = \mathbf{0}$, $\mathbf{C}_0^{r(m)} = \mathbf{0}$.
$\mathbf{\Delta}_0^{(m)^x} = \mathbf{0}$, $\mathbf{\Delta}(m)_0^x = \mathbf{0}$.

The output of the filter is

$$z_i(n) = \tilde{z}_i(n) - e_n^{(L_2,i)^x}.$$

The recursive inversion of the matrices $\mathbf{C}_n^{e(m)}$, $\mathbf{C}_n^{r(m)}$ can introduce numerical difficulties. We employ the solution proposed by Lewis in [14] who formulated the typical lattice recursions in terms of QR decompositions of the forward and the backward error matrices. This is achieved by a factorization of the matrices $\mathbf{C}_n^{e(m)}$, $\mathbf{C}_n^{r(m)}$, observing that they have a structure similar to a correlation matrix (i.e., $Z^{T^*}Z$). Using QR decomposition of the matrices resulting from the factorization of $\mathbf{C}_n^{e(m)}$, $\mathbf{C}_n^{r(m)}$, $\mathbf{\Delta}_n^{(m)}$, it is possible to express the recursions as in the following [14]

$$\begin{bmatrix} (\gamma_{n-1}^{(m-1)})^{-1}\mathbf{e}_n^{(m-1)^T} & (\gamma_{n-1}^{(m-1)})^{-1}\mathbf{r}_{n-1}^{(m-1)^T} & \gamma_{n-1}^{(m-1)} \\ \lambda\mathbf{R}_{n-1}^{e(m)} & \lambda\mathbf{X}_{n-1}^{e(m)} & \mathbf{0} \end{bmatrix}$$

$$\rightarrow (\text{Givens Rotations}) \rightarrow \begin{bmatrix} \mathbf{0}^T & \times & \times \\ \mathbf{R}_n^{e(m)} & \mathbf{X}_n^{e(m)} & \beta_n^{e(m)} \end{bmatrix}$$

$$\begin{bmatrix} (\gamma_{n-1}^{(m-1)})^{-1}\mathbf{r}_{n-1}^{(m-1)^T} & (\gamma_{n-1}^{(m-1)})^{-1}\mathbf{e}_n^{(m-1)^T} & \gamma_{n-1}^{(m-1)} & (\gamma_{n-1}^{(m-1)})^{-1}\mathbf{e}_n^{(m-1)^{xT}} \\ \lambda\mathbf{R}_{n-1}^{r(m)} & \lambda\mathbf{X}_{n-1}^{r(m)} & \mathbf{0} & \lambda\mathbf{X}_{n-1}^{x(m)} \end{bmatrix}$$

$$\rightarrow (\text{Givens Rotations}) \rightarrow \begin{bmatrix} \mathbf{0}^T & \times & \times & \times \\ \mathbf{R}_n^{r(m)} & \mathbf{X}_n^{r(m)} & \beta_n^{r(m)} & \mathbf{X}_n^{x(m)} \end{bmatrix}$$

$$\mathbf{r}_n^{(m)} = \mathbf{r}_{n-1}^{(m-1)} - \mathbf{X}_n^{e(m)^{T^*}}\beta_n^{e(m)}$$

$$\mathbf{e}_n^{(m)} = \mathbf{e}_n^{(m-1)} - \mathbf{X}_n^{r(m)^{T^*}}\beta_n^{r(m)}$$

$$\mathbf{e}_n^{(m)^x} = \mathbf{e}_n^{(m-1)^x} - \mathbf{X}_n^{x(m)^{T^*}}\beta_n^{r(m)}$$

$$(\gamma_n^{(m)})^2 = (\gamma_{n-1}^{(m-1)})^2 - \beta_n^{e(m)^{T^*}}\beta_n^{e(m)}.$$

The initialization is

$$\mathbf{R}_n^{r(m)} = \mathbf{X}_n^{r(m)} = \mathbf{R}_n^{e(m)} = \mathbf{X}_n^{e(m)} = \mathbf{X}_n^{x(m)} = \mathbf{0}.$$

It is well known that the improved numerical behavior of square-root algorithms is due in large part to the reduction of the numerical range of the variables. Generally, if the inversion update for the autocorrelation matrix is implemented using the traditional RLS-Kalman update one needs twice the numerical precision to achieve the same accuracy of the proposed QR-lattice algorithm [14, 18].

VI. RESULTS OF SIMULATIONS

It was shown [15] that for mobile satellite communications (VHF/UHF low Earth orbiting) the RF link experiences propagation problems mainly in the form of shadowing by trees, buildings, etc. Still the vegetative and structure shadowing is the main impairment in the L and S bands GSO mobile communications. At higher frequencies atmospheric effects gain importance, and while if at C-band these effects are not severe (that is the main reason for the great success of C-band geostationary satellites) current developments in research programs indicate the high interest in frequencies above 10 GHz. For example, Ka-band GSO fixed or mobile communications (VSAT systems), but also GSO communications above 30 GHz suffer by gaseous loss, cloud loss, rain fade, and rain depolarization.

We performed simulations of the adaptive method using a word length size of 24 bits, using fixed-point arithmetic. Digital signal processing processors are commercially available with this characteristics. We also reduced the number of bits to 16 to use fixed-point arithmetic processors with different costs without observing major degradations of performance. As expected, the QR-lattice filter exhibited excellent numerical behavior and robustness to roundoff errors. The traditional Kalman recursive least squares algorithm required instead double precision to achieve reasonable convergence in any channel condition.

The simulations are intended to study QPSK in presence of multipath and CPI. CNR and CIR are defined as CNR = $(P_{av}/2)/(N_0/T)|_{\text{dB}}$ with $P_{av} = P_v + P_h$, P_v and P_h as defined in Section II,

$$\text{CIR} = \left.\frac{\int |\mathcal{F}_v(\omega)|^2\, d\omega}{\int |\mathcal{F}_{h,v}(\omega)|^2\, d\omega}\right|_{\text{dB}} = \left.\frac{\int |\mathcal{F}_h(\omega)|^2\, d\omega}{\int |\mathcal{F}_{v,h}(\omega)|^2\, d\omega}\right|_{\text{dB}}.$$

The multipath model for $f_v(t)$ and $f_h(t)$ is a two-ray Rayleigh fading model. This means that $N_v = N_h = 2$, while $\rho_{v,m}$, $\rho_{h,m}$ for $m = 1,2$ are four independent complex zero-mean Gaussian processes with variances $\sigma_{v,m}^2$, $\sigma_{h,m}^2$, defined by the desired delay spread profile. We assume that the two rays in the main channels and the 2 independent rays in the cross-channels have the same power (which is a worst case scenario reflecting severe distortion), and that the impulse response of the multipath channel is normalized to unit energy. The delays $\tau_{v,2}$, $\tau_{h,2}$ are fixed as specified in every experiment and normalized to the symbol period, while it is assumed that $\tau_{v,1} = \tau_{h,1} = 0$.

The Doppler frequency usually describes the SOS of channel variations. The model used is based on the wide sense stationary uncorrelated scattering (WSSUS) assumption [12], so that $\rho_{v,m}$, $\rho_{h,m}$, R_m for $m = 1,2$ are generated as filtered Gaussian processes (with variances $\sigma_{v,m}^2$, $\sigma_{h,m}^2$, $\sigma_{R_1}^2$, $\sigma_{R_2}^2$, respectively) fully specified by the scattering function. Particularly each process has a frequency response equal to the square-root of the Doppler power density spectrum. We approximate the Doppler spectrum with rational filtered processes. The filters are described by their 3 dB bandwidth which is called the normalized Doppler frequency. The additional assumption is that all the complex factors on the vertically and horizontally polarized channel have the same Doppler spectrum.

The power of the interference is totally controlled by K_1, K_2, K_3, K_4, $\sigma_{R_1}^2$, $\sigma_{R_2}^2$.

The model is not intended to depict the underlying physical mechanism of propagation effects on satellite communication links for a particular application. However, it is general because it can characterize phase and amplitude distortions (due to multipath propagation) of the main channels (vertical and horizontal) as well the CPI resulting from RF propagation and represents a good assumption in evaluating the technique we are proposing.

We also investigated the performance of the algorithm by comparing it with the second-order statistic-based linear MMSE filter implemented as a RLS Kalman filter and the simple gradient-based blind methods proposed in [16] based on the CMA and in [17] based on the Sato algorithm (actually, its generalization to the complex case).

The RLS Kalman estimator is a filter having perfect knowledge of the transmitted sequence [12, 18]. This multichannel adaptive filter is updated using the RLS algorithm of [14].

The CMA approach is based on a popular stochastic gradient-based search idea proposed by Godard [16]. The update of the coefficients is

$$\tilde{\mathbf{w}}_i(k+1) = \tilde{\mathbf{w}}_i(k) + \mu \tilde{\mathbf{y}}^*(k) z_i(k) |z_i(k)|^{\tilde{p}-2} (|z_i(k)|^{\tilde{p}} - R_{\tilde{p}})$$

where $\tilde{\mathbf{w}}_i(k)$ is the vector of filter weights at time step k as it is defined in (7), $\tilde{\mathbf{y}}(k)$ is defined in (22),

$$R_{\tilde{p}} = \frac{E\{|x_i(n)|^{2\tilde{p}}\}}{E\{|x_i(n)|^{\tilde{p}}\}}$$

and μ is the adaption gain. The CMA method is obtained for $\tilde{p} = 2$.

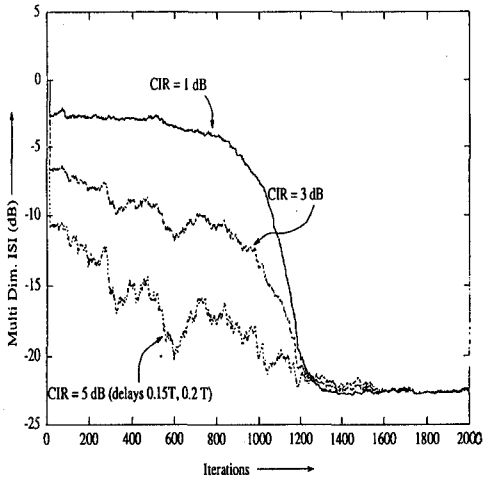


Fig. 3. Results for $L = 7$, CNR = 10 dB and at different CIR values. Channel is static in each run but has different realization from run to run. Average is computed over 100 Monte Carlo runs. Delays are $\tau_{v,1} = \tau_{h,1} = 0$, $\tau_{v,2} = \tau_{h,2} = 0.25$, $D_1 = D_2 = 0.15$.

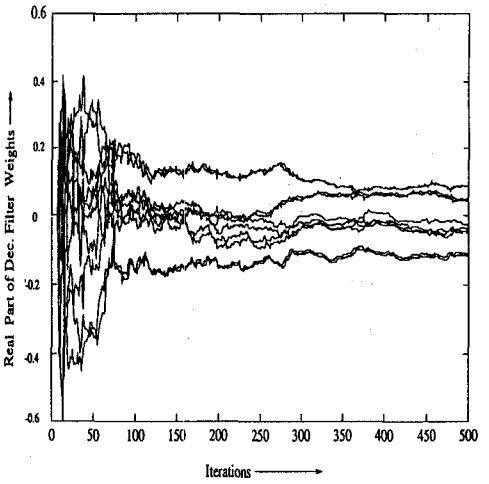


Fig. 4. Convergence event for real part of coefficients representing impulse response of multichannel lattice, under same conditions of Fig. 3 and CIR = 5 dB.

The Sato algorithm uses the same multichannel architecture and the same gradient-based procedure, but the nonlinear cost function is given by the (complex) sign function so that the update algorithm is described by

$$\tilde{\mathbf{w}}_i(k+1) = \tilde{\mathbf{w}}_i(k) + \mu\tilde{\mathbf{y}}^*(k)[z_i(k) - c\,\mathrm{sign}(z_i(k))]$$

where $c\,\mathrm{sign}(x) = \mathrm{sign}(\mathrm{Re}\{x\}) + j\,\mathrm{sign}(\mathrm{Im}\{x\})$.

We evaluate the instantaneous performance of the system for the vertically polarized channel as the multichannel ISI (intersymbol interference)

$$MDI_v(\mathbf{S}) = MDI_1(\mathbf{S}) = \frac{\sum_{j=1}^{2}\sum_l |s_{1,j}(l)|^2 - |s_{1,j}^{\max}|^2}{|s_{1,j}^{\max}|^2}$$

where $s_{i,j}^{\max} = \max_{k=\ldots,-1,0,1,\ldots;j=1,2}[s_{i,j}(k)]$.

In Fig. 3 we evaluate the case when the delays for the reflected paths are $\tau_{v,2} = \tau_{h,2} = 0.25$, $D_1 = D_2 = 0.15$ at different CIRs with CNR = 10 dB and $L = 7$. The channel is static in each run but the $\mathrm{MDI}_1(\mathbf{S})$ is averaged over 100 Monte Carlo runs, and each experiment presents a different realization of the channel parameters.

For CNR = 15 dB, CIR = 5 dB and same situation of Fig. 3, the convergence rate of the real part of the restoration filter impulse response is shown in Fig. 4.

Fig. 5 shows (in the same channel conditions of Fig. 3 for CIR = 10 dB) the above mentioned comparison with the CMA algorithm with the lattice fourth-order cumulant-based method here presented. The improvement in speed is evident and is mainly caused by the efficient orthogonalization performed by the lattice section of the filter.

Fig. 6 (same situation of Fig. 5) reports the performance of the proposed method with respect to the RLS Kalman filter, whose error function benefits

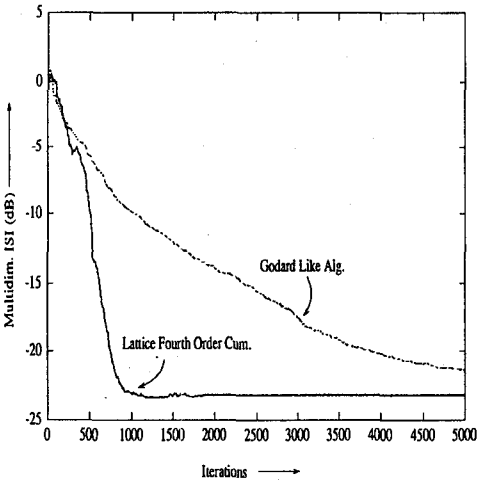


Fig. 5. Convergence of proposed HOS-lattice filter (based on fourth-order cumulants) as opposed to performance of CMA (Godard-like) method under same conditions of Fig. 3, for CIR = 10 dB.

of a perfect knowledge of the transmitted signal. Also the Sato algorithm performance is shown. The proposed blind method is considerably slower than a trained method. The improved identification properties consequent to the use of high-order statistics are evident by a lower ISI (the RLS algorithm remains flat at about −20 dB, while the lattice-HOS filter continues to improve as new samples are processed).

In Fig. 7 we show the bit error rate (BER) results of the simulations versus CNR for a channel with delays in the reflected paths $\tau_{v,2} = \tau_{h,2} = 0.15$ and $D_1 = D_2 = 0.05$ for different values of CIR and with a deconvolution filter length equal to $L = 3$. BER measurement is performed after the first 500 symbols, which exclude the transient behavior of the initial convergence process. In the BER evaluation the channel is time varying with the product of the

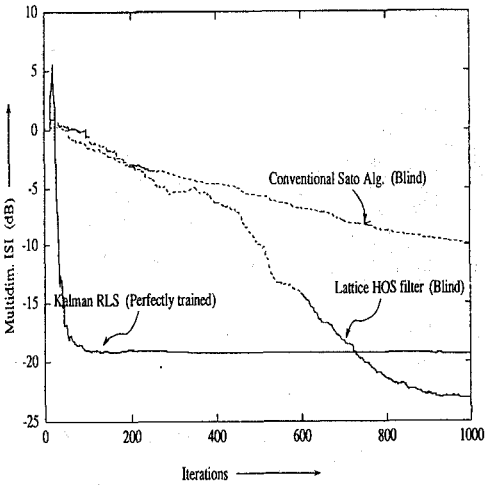


Fig. 6. Convergence of proposed HOS-lattice filter (based on fourth-order cumulants) as opposed to performance of Kalman-RLS method and SATO algorithm (same channel conditions of Fig. 5). RLS algorithm is not blind. It is assumed to have complete knowledge of transmitted symbols.

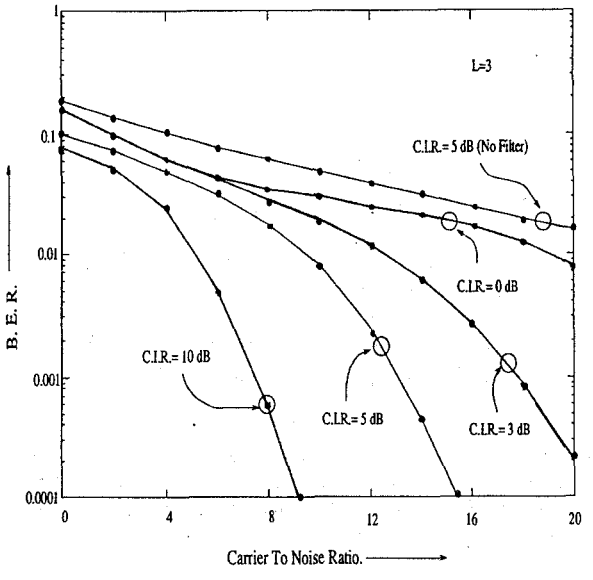


Fig. 7. BER for time-varying channel with $f_DT = 0.0004$ at different levels of CPI. Delays are $\tau_{v,1} = \tau_{h,1} = 0$, $\tau_{v,2} = \tau_{h,2} = 0.15$, $D_1 = D_2 = 0.05$. Filter has length $L = 3$.

maximum Doppler frequency f_D and the symbol period T given by $f_DT = 0.0004$. Also the BER performance of the system without compensation filter (labeled *no filter*) is shown for CIR = 5 dB.

In Fig. 8 BER results are shown for severe cochannel and multipath distortions ($\tau_{v,2} = \tau_{h,2} = 0.35$, $D_1 = D_2 = 0.25T$, $f_DT = 0.0003$), and using the same methodology used in Fig. 7. The implicit trade off between performance and complexity required for the choice of the length of the restoration filter L is evident.

Fig. 9 shows tracking performance of the algorithm over a time-varying channel when $\tau_{v,2} =$

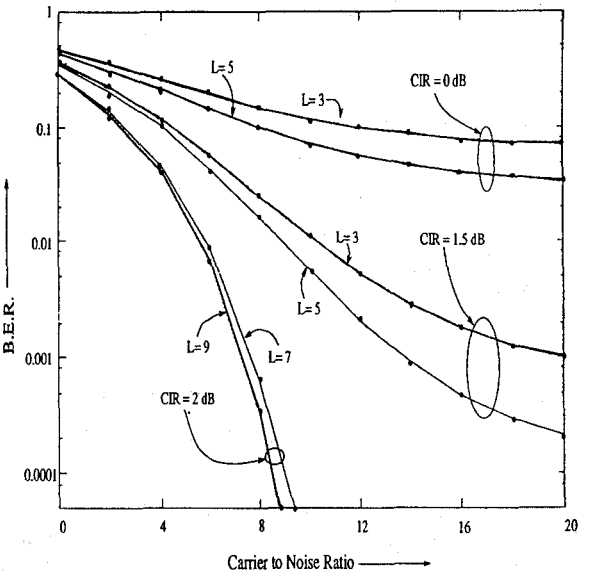


Fig. 8. BER for time-varying channel with $f_DT = 0.0003$ at different level of PPI and at different level of hardware complexity (different L values) with severe multipath. The delays are $\tau_{v,1} = \tau_{h,1} = 0$, $\tau_{v,2} = \tau_{h,2} = 0.35$, $D_1 = D_2 = 0.25$.

$\tau_{h,2} = 0.3$ and $D_1 = D_2 = 0.1$, $f_DT = 0.0006$. The real part of the restoration filter impulse response (which in our implementation is a lattice filter) is plotted against the optimum asymptotic performance obtained by computation of (9) as the channel changes.

A sudden rain event, that is a sudden change in the propagation characteristics, is also simulated by a sudden decrease of the CIR of the matrix transfer function $C(\omega)$ from 20 dB to 5 dB when CNR = 10 dB with $\tau_{v,2} = \tau_{h,2} = 0.2$ and $D_1 = D_2 = 0.05$. The results of this experiment are reported in Fig. 10.

VII. CONCLUSIONS

We have presented a blind adaptive multichannel linear filter for compensation of CPI and multipath propagation for OBP satellites. The proposed algorithm is particularly attractive due to computational efficiency. The high performance of the high-order statistics-based estimation approach is achieved by means of a lattice architecture. Considerable improvement in convergence rate with respect to other blind identification procedures is possible due to a specific implementation of the orthogonalization of the received samples. The results of the computer simulations and a theoretical analysis of the residual error at the point of convergence confirm the effectiveness of the method.

APPENDIX A. PERFORMANCE ANALYSIS

Some of the results derived in [4] can be used to analyze the algorithm in order to obtain asymptotic

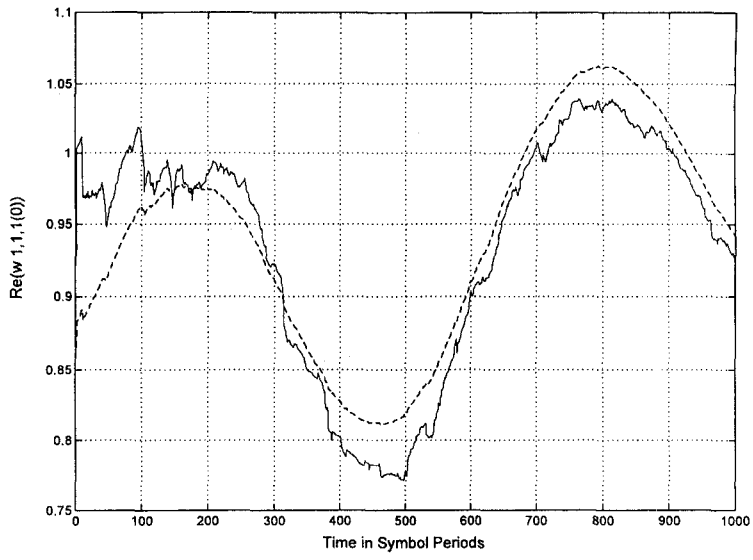


Fig. 9. Tracking performance of HOS-lattice approach with CNR 30 dB. Channel is varying and product maximum Doppler frequency-symbol period is equal to 0.0006. Delays are $\tau_{v,1} = \tau_{h,1} = 0$, $\tau_{v,2} = \tau_{h,2} = 0.3$, $D_1 = D_2 = 0.1$. Dashed curve is trace of real part of weight $W_{1,1}(0)$ obtained computing (9) while solid curve is obtained using adaptive algorithm.

performance evaluation. We outline a procedure to obtain an error rate bound that can be used to take important system level decisions on the actual parameters of the digital signal processing section of the receiver (for example filter lengths). The following relation holds (we indicate estimated quantities as $\overline{(\cdot\cdot)}$):

$$\overline{\mathbf{R}}_{m,n} = \sum_{j_1=1}^{2}\sum_{j_2=1}^{2} \mathbf{H}_{m,j_1}^{T^*}\overline{\mathbf{A}}_{m,n}\mathbf{H}_{n,j_2}$$

where the n,m element of matrices $\overline{\mathbf{A}}_{i,j}$ is given by

$$[\overline{\mathbf{A}}_{i,j}]_{l,m} = \frac{\overline{\mathrm{cum}}[x_i(n-l), x_j^*(n-m)]}{\sigma_x^2}.$$

Now consider the matrix $\overline{\overline{\mathbf{R}}}$, that is the matrix obtained by the cumulants estimation based on sample averages, and the approximation

$$\overline{\overline{\mathbf{R}}}^{-1} \simeq \tilde{\mathbf{R}}^{-1} - \tilde{\mathbf{R}}^{-1}(\overline{\overline{\mathbf{R}}} - \tilde{\mathbf{R}}^{-1}))\tilde{\mathbf{R}}^{-1}$$

$$= (\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1} - (\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}(\overline{\tilde{\mathbf{A}}} - \mathbf{I}_{2L})\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}$$

where $\overline{\tilde{\mathbf{A}}}$ is the block matrix with the generic i,j block given by $\overline{\mathbf{A}}_{i,j}$. The iterative procedure (10)–(11) applied to the vector $\tilde{\mathbf{s}}_i$ maintains the index of the tap with largest magnitude which was called in [4] the *leading tap* (see, for details, [4, Section III]). Assuming the reference channel is the vertical channel (channel 1), and that the *leading tap* is $s_{1,1}(0)$ observe that $\overline{\mathbf{D}}_1 \simeq \tilde{\mathbf{H}}^{T^*}\overline{\overline{\mathbf{g}}}_1$ where $\overline{\overline{\mathbf{g}}}_1 = (\overline{\mathbf{g}}_{1,1}^T, \overline{\mathbf{g}}_{1,2}^T, \overline{\mathbf{g}}_{1,3}^T, \ldots, \overline{\mathbf{g}}_{1,2}^T)^T$

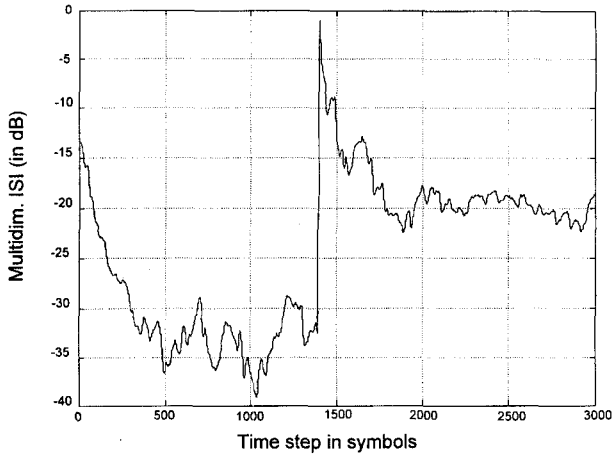


Fig. 10. Response of adaptive filter to a sudden change of the propagation conditions (for example a sudden rainfall). Delays are $\tau_{v,1} = \tau_{h,1} = 0$, $\tau_{v,2} = \tau_{h,2} = 0.2$, $D_1 = D_2 = 0.05$, CNR = 10 dB, CIR is changed from 20 dB to 5 dB.

and the nth element of the vector $\overline{\mathbf{g}}_{i,j}$ is

$$[\overline{\mathbf{g}}_{i,j}]_n = s_{1,1}^2(0)s_{1,1}^*(0)\frac{\overline{\mathrm{cum}}[x_i(t); x_i(t); x_i^*(t); x_j^*(t-n)]}{\mathrm{cum}_4(x)}.$$

If L (length of the deconvolution filter) is taken large enough, $\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}$ will tend towards the identity matrix and $\overline{\overline{\mathbf{g}}}_1 \simeq s_{1,1}(0)^2 s_{1,1}^*(0)\tilde{\delta}_1$. Thus

$$\tilde{\mathbf{w}}_1^{[1]} = \overline{\overline{\mathbf{R}}}^{-1}\overline{\mathbf{D}}_1 \simeq (\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\mathbf{u}_1$$

with

$$\mathbf{u}_1 = \overline{\overline{\mathbf{g}}}_1 - (\overline{\tilde{\mathbf{A}}} - \mathbf{I}_{2L})\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\overline{\overline{\mathbf{g}}}_1$$

$$\simeq s_{1,1}(0)^2 s_{1,1}^*(0)[\tilde{\delta}_1 + (\mathbf{I}_{2L} - \overline{\tilde{\mathbf{A}}})\tilde{\delta}_1].$$

We can now write the predicted result for the impulse response of the deconvolution filter as

$$\tilde{\mathbf{w}}_1^{[2]} = \frac{\tilde{\mathbf{w}}_1^{[1]}}{\sqrt{\tilde{\mathbf{w}}_1^{[1]T}\tilde{\bar{\mathbf{R}}}\tilde{\mathbf{w}}_1^{[1]}}} = \frac{(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\mathbf{u}_1}{\sqrt{\mathbf{u}_1^{T^*}\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\bar{\mathbf{A}}}\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\mathbf{u}_1}} \simeq \frac{(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\mathbf{u}_1}{\sqrt{\mathbf{u}_1^{T^*}\tilde{\bar{\mathbf{A}}}\mathbf{u}_1}}. \quad (24)$$

The decision variable for the nth bit of the vertical or horizontal channel can be expressed in a scalar form as

$$z_i(n) = s_{i,i}(0)x_i(n) + \sum_j' \sum_k' s_{i,j}(n-k)x_j(k) + \sum_{k=L_1}^{L_2}\sum_{j=1}^{2} w_{i,j}(n-k)n_j(k) \qquad i = 1,2 \quad (25)$$

where the double sum $\sum_j' \sum_k'$ excludes the term with $j = i$, $k = n$. A simplifying assumption commonly made, is that one can focus on the *nonexcess bandwidth transmission* case, that is we assume that $\mathcal{P}_{rx}(\omega)\mathcal{C}(\omega)\mathcal{P}_{tx}(\omega) = 0$ for $|\omega| \geq \pi/T$ so that the noise power is measured in the Nyquist band (the receiver filter is perfectly bandlimited). The term $\gamma_i(n) = \sum_{k=L_1}^{L_2}\sum_{j=1}^{2} w_{i,j}(n-k)n_j(k)$ in (25) has Gaussian distribution with zero mean and variance

$$\sigma_n^2 = E\{|\gamma_i(n)|^2\} = \sum_{j_1,j_2=1}^{2}\sum_{k_1,k_2=L_1}^{L_2} w_{i,j_1}(k_1)w_{i,j_2}^*(k_2) \times E\{n_{j_1}(n-k_1)n_{j_2}^*(n-k_2)\} = N_0\sum_{j=1}^{2}\sum_{k=L_1}^{L_2} w_{i,j}(k)w_{i,j}^*(k).$$

In the case of QPSK the decisions with thresholds equal to 0 are taken independently on $\mathrm{Re}[z_i(n)]$ and $\mathrm{Im}[z_i(n)]$ (the method can be also extended to any digital modulation with rectangular constellation). These two random variables have Gaussian distribution, variance given by $E\{\mathrm{Re}[\gamma_i(n)]^2\} = E\{\mathrm{Im}[\gamma_i(n)]^2\} = \frac{1}{2}\sigma_n^2$ and mean easily derived from (25). For a fixed sequence of symbols transmitted on the vertical and horizontal channel the probability of making an error in the decision on $\mathrm{Re}[z_i(n)]$ and $\mathrm{Im}[z_i(n)]$ can be expressed, due to the Gaussian noise term, as an erf(·) function related to the variance of the noise term. Then one should compute the expected value of this quantity with respect to all the possible transmitted sequences, using the probability density function of the term $\sum_j' \sum_k' s_{i,j}(n-k)x_j(k)$ in (25), which is analogous to when one tries to compute the probability of error in a communication system afflicted by ISI [12, 19]. Unfortunately it is well known that this is a problem of considerable complexity and several approaches were proposed over the years to approximate the exact probability of error in this case.

One of the methods to evaluate a bound on the error rate of a system with ISI is the Saltzberg method [19] as used in [6], based on Chernoff bounding techniques. This immediately gives for the probability of error on the vertical channel and the horizontal channel

$$P_e^{(v)} \leq \exp\left(-\frac{1}{2}\frac{|s_{1,1}(0)|^2}{\sigma_n^2/2 + \sigma_x^2/2\sum_j''\sum_k''|s_{1,j}(k)|^2}\right) \quad (26)$$

$$P_e^{(h)} \leq \exp\left(-\frac{1}{2}\frac{|s_{2,2}(0)|^2}{\sigma_n^2/2 + \sigma_x^2/2\sum_j''\sum_k''|s_{2,j}(k)|^2}\right) \quad (27)$$

where the summations $\sum_j''\sum_k''$ are summations over $j = 1,2$ and $k = \ldots,-1,0,+1,\ldots$ which exclude the term $s_{1,1}(0)$ in (26) and the term $s_{2,2}(0)$ in (27). It is important to observe that the $s_{i,j}(k)$ response needs to be computed from $\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i$ for $i = 1$ (the vertical channel) and $i = 2$ (the horizontal channel), using for $\tilde{\mathbf{w}}_i$ the asymptotic result (24). The bounds (26), (27) are not very tight, but, as already observed, they can be successfully used as an indication of the performance related to the important parameters of the system.

Obviously the derived expressions are related to a particular realization of the multipath channels. The result should be then averaged with respect to the time-varying characteristics of the fading processes.

APPENDIX B. ON THE CONVERGENCE OF THE ALGORITHM

From $\tilde{\mathbf{s}}_i^0$ with elements $s_{i,j}^0(k)$ we want to achieve convergence after l iterations to $\tilde{\mathbf{s}}_i^l$ with elements $s_{i,j}^l(k)$ that are non-zero only for a specific $j = j_0$ and $k = k_0$. While there is no specific requirement on k_0 (which will result in an arbitrary delay) it is evident that it must hold $j_0 = i$ to restore the information of the ith channel and not an arbitrary j_0th channel. This is assured if the leading tap of $\tilde{\mathbf{s}}_i^0$ is contained in the vector $\mathbf{s}_{i,i}^0$, that is if $|s_{i,i}^0(k_0)| > |s_{i,j}^0(k)|$ is verified for some k_0, all $j \neq i$ and any k. This situation is always verified as long as one initializes $\tilde{\mathbf{w}}_i$ with $\mathbf{w}_{i,j} = 0$ for all $j \neq i$. The initialization of $\mathbf{w}_{i,i}$ is arbitrary. Another important aspect is the global convergence of the algorithm in the $\tilde{\mathbf{w}}_i$ domain to the same solution of the algorithm in the $\tilde{\mathbf{s}}_i$ domain: this cannot be generally guaranteed when $\tilde{\mathbf{w}}_i$ has finite length. The algorithm

in the $\tilde{\mathbf{w}}_i$ domain (13)–(14) *projected* back in the $\tilde{\mathbf{s}}_i$ domain becomes

$$\tilde{\mathbf{s}}_i^{[1]} = \tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{g}}_i \qquad \tilde{\mathbf{s}}_i^{[2]} = \frac{\tilde{\mathbf{s}}_i^{[1]}}{\|\tilde{\mathbf{s}}_i^{[1]}\|}$$

whose point of convergence easily obtained (see [4]) is

$$\tilde{\mathbf{s}}_i \simeq \frac{1}{\sqrt{\tilde{\delta}_i^T\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\delta}_i}}\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\delta}_i.$$

This expression is coincident with $\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}(\tilde{\mathbf{H}}^{T^*}\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^{T^*}\tilde{\delta}_i$ solution of $\min_{\tilde{\mathbf{w}}_i}\|\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i - \tilde{\delta}_i\|^2$, up to a gain factor. The problem of the convergence of a certain function $\Psi_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i) = \Psi(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i)$ (when $\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i$) to the same solution of the function $\Psi(\tilde{\mathbf{s}}_i)$ is a well known and investigated problem in the blind deconvolution literature (see [20, chapter by Shalvi and Weinstein]). This convergence cannot always be guaranteed for both functions. Shalvi and Weinstein show in [20, Appendix E, ch. 5], the conditions under which this is true for the impulse response to be deconvolved. Those considerations can be applied to the super exponential algorithm, observing that procedure (10)–(11) is a gradient-based search to solve the maximization problem

$$\max_{\tilde{\mathbf{s}}_i}\Psi(\tilde{\mathbf{s}}_i) \tag{28}$$

subject to the constraint $\|\tilde{\mathbf{s}}_i\|^2 = 1$. In fact the two steps in (10)–(11) are equivalent to the gradient-based iteration

$$\tilde{\mathbf{s}}_i = \tilde{\mathbf{s}}_i + \delta\frac{\partial\Psi(\tilde{\mathbf{s}}_i)}{\partial\tilde{\mathbf{s}}_i}$$

with $\Psi(\tilde{\mathbf{s}}_i) = \sum_j\sum_k s_{i,j}(k)^2 s_{i,j}(k)^{*2}$, $\partial f(\mathbf{v})/\partial\mathbf{v}$ indicates the gradient of the vector $\mathbf{v}$ and we choose a *very large* step size δ. When we maximize $\Psi_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{H}}\mathbf{w}_i)$, generally we do not converge to the same solution of (28). In [20] it is demonstrated that the extremum that we find working on $\tilde{\mathbf{w}}_i$ is arbitrarily close to the extremum of (28), if the system $\tilde{\mathcal{H}}(z)$ is stable (this condition is in our model assumptions) and if the order of $\tilde{\mathbf{w}}_i$ is the one that allows *satisfactory* deconvolution (in the case of deconvolution of a finite-impulse response (FIR) system, perfect deconvolution with a linear FIR model cannot be achieved). Important results for this class of algorithms can also be found in [21].

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Prof. Gary J. Saulnier and Prof. John B. Anderson, ECSE Department, Rensselaer Polytechnic Institute, for their support and their suggestions.

REFERENCES

[1] Chu, T. S. (1973)
Restoring the orthogonality of two polarizations in radio communications systems, I.
Bell System Technical Journal, **50**, 9 (Mar. 1973), 3063–3069.

[2] Nichols, H. E., Giordano, A. A., and Proakis, J. G. (1977)
MLD and MSE algorithms for adaptive detection of digital signals in the presence of interchannel interference.
IEEE Transactions on Information Theory, **IT-23**, 5 (Sept. 1977), 563–575.

[3] Swami, A., Giannakis, G., and Shamshnder, S. (1994)
Multichannel ARMA processes.
IEEE Transactions on Signal Processing, **42** (Apr. 1994), 898–913.

[4] Shalvi, O., and Weinstein, E. (1993)
Super exponential methods for blind deconvolution.
IEEE Transactions on Information Theory, **39** (Mar. 1993).

[5] Bello, P. A. (1963)
Charachterization of randomly time variant linear channels.
IEEE Transactions on Communication Technology, **COM-11** (Dec. 1963), 360–393.

[6] Kahverad, M., and Salz, J. (1985)
Cross-polarization cancellation and equalization in digital transmission over dually polarized multipath fading channels.
AT&T Technology Journal, **64**, 10 (Dec. 1985), 2211–2245.

[7] Amitay, N., and Salz, J. (1984)
Linear equalization theory in digital data transmission over dually polarized fading radio channels.
AT&T Bell Laboratory Technology Journal, **63**, 10 (Dec. 1984), 2215–2259.

[8] Mendel, J. M. (1991)
Tutorial on higher order statistics (spectra) in signal processing and system theory: Theoretical results and some applications.
Proceedings of the IEEE, **79** (Mar. 1991), 278–305.

[9] Tong, L., Liu, R., Soon, V., and Huang, Y. (1991)
Indeterminacy and identifiability of blind identification.
IEEE Transactions on Circuits and Systems, **38** (May 1991), 499–509.

[10] Li, Y., and Liu, K. J. R.
Adaptive blind multi-channel equalization for multiple signals separation.
IEEE Transactions on Signal Processing, submitted for publication.

[11] Satorius, E. H., and Pack, J. D. (1981)
Applications of least squares lattice algorithms to adaptive equalization.
IEEE Transactions on Communication, **COM-29** (Feb. 1981), 136–142.

[12] Proakis, J. G. (1989)
Digital Communications.
New York: McGraw-Hill, 1989.

[13] Cioffi, J. M., and Kailath, T. (1984)
Fast recursive least squares transversal filters for adaptive filtering.
IEEE Transactions on Acoustics, Speech, and Signal Processing, **ASSP-32** (Apr. 1984), 302–337.

[14] Lewis, P. (1990)
QR based algorithms for multichannel adaptive least squares lattice filters.
IEEE Transactions on Acoustics, Speech, and Signal Processing, **30** (Mar. 1990), 421–432.

[15] Stutzman, W. L. (1993)
Prolog on the special section on propagation effects on satellite communication links.
Proceedings of the IEEE, **81**, 6 (June 1993), 850–855.

[16] Godard, D. N. (1980)
Self-recovering equalization and carrier tracking in two dimensional data communications systems.
IEEE Transactions on Communication, **COM-28** (Nov. 1980), 1867–1875.

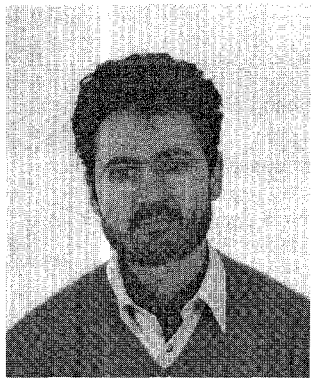
[17] Sato, Y. (1975)
A method of self-recovering equalization for multi-level amplitude-modulation systems.
IEEE Transactions on Communication, **COM-23** (June 1975), 679–682.

[18] Haykin, S. (1986)
Adaptive Filter Theory.
Englewood Cliff, NJ: Prentice Hall, 1986.

[19] Saltzberg, B. R. (1968)
Intersymbol interference error bounds with application to ideal bandlimited signaling.
IEEE Transactions on Information Theory, **IT-14** (July 1968), 563–568.

[20] Shalvi, O., and Weinstein, E. (1994)
Universal methods for blind deconvolution.
In S. Haykin (Ed.), *Blind Deconvolution*.
Englewood Cliffs, NJ: Prentice-Hall, 1994.

[21] Li, Y., and Ding, Z. (1995)
Convergence analysis of finite length blind adaptive equalizers.
IEEE Transactions on Signal Processing, **43** (Sept. 1995), 2120–2129.



Massimiliano (Max) Martone (M'93) was born in Rome, Italy. He received the degree of Doctor in electronic engineering in 1990 at the University of Rome "La Sapienza."

He worked for the Italian Air Force in 1990–1991 and consulted in the area of digital signal processing (DSP) applied to communications for Staer Inc., S.P.E. Inc., TRS-Alfa Consult Inc. In 1991 he joined the DSP staff of the On Board Equipment Division of Alenia Spazio where he was involved in the design of DSP-based receivers and spread spectrum TT&C transponders. He also collaborated with the digital communications research group of Fondazione Ugo Bordoni. In 1994 he was appointed Visiting Scientist at the ECSE Department of Rensselaer Polytechnic Institute, Troy, NY. He was a wireless communications consultant for ATS Inc., Waltham, MA, and in 1995 he joined the Telecommunications Group of Watkins-Johnson Company, Gaithersburg, MD. He is also with the George Washington University, Electrical Engineering Department. His main interests are in advanced signal processing for HF receivers implementation, spread spectrum multiple access communications, mobile radio communications.

Dr. Martone is a member of the New York Academy of Sciences and the Associazione Elettrotecnica Italiana.