Cumulant-Based Adaptive Multichannel Filtering for Wireless Communication Systems with Multipath RF Propagation Using Antenna Arrays

Massimiliano (Max) Martone, *Member, IEEE*

***Abstract*—A method based on high-order statistics is proposed to mitigate the performance degradation caused by multipath RF propagation in a mobile radio communication system using a linear antenna array at the base-station receiver. It is shown that an overdetermined system of linear equations (involving only cumulants of the received baseband digitized signal) can be obtained to perform noniterative deconvolution. An efficient adaptive algorithm based on square-root decomposition is proposed to avoid numerical problems when real-time tracking of moving transmitters is needed.**

I. Introduction

THE USE of antenna arrays in wireless communications can theoretically improve system performance in terms of signal quality and capacity. Particularly, a multielement antenna at the base-station receiver of a cellular system is able to compensate signal degradations in the mobile to base link [12], [13]. It is important to observe, however, that a practical and definitive solution for digital cellular systems employing time-division multiple access (TDMA) does not exist at this time.

In TDMA systems, data dispersion can span several symbols as a consequence of frequency-selective fading caused by RF multipath propagation. In addition, propagation characteristics may change in time due to the motion of the transmitter. The received signal is composed of the original plus several delayed attenuated replicas, and each replica reaches the antenna with a different attenuation and angle of arrival. Space-only processing methods are not effective because intersymbol interference (ISI) cannot be compensated using the traditional beamforming architecture. A careful combination of space and time filtering may result in an extremely efficient method to solve ISI caused by multipath fading. This subject was studied in great detail in [1] and [11], where it was shown that the use of joint diversity/equalization methods offers significant advantages with respect to traditional systems. The structure we propose and analyze is different from the optimum scheme since a digital filter is provided for each diversity branch. This architecture was proposed by antenna arrays designers and successfully applied on radars more than a decade ago as a method to increase bandwidth resolution [7]. The implementation was entirely based on analog circuitry (delay lines and weights control): this made the idea not very attractive in terms of cost and reliability. Recent progress in digital technology provides solutions to difficult implementation problems caused by high-complexity algorithms for adaptive antennas.

The space-time filtering approach used in this paper results in a discrete-time single-input–multiple-outputs (SIMO's) model that must be deconvolved (equalized). Standard techniques for equalization are based on an equivalent minimum phase (MP) system modeling approach because they exploit only the second-order statistics (SOS's) of the signal. The minimum-mean-square-error (MMSE) criterion, for example, is based on SOS only. However, the most part of all real-world channels do not present the MP condition. Motivated by these observations, the equalization of nonminimum phase (NMP) channels using higher than second-order statistics (HOS) has stimulated incredible interest during the last ten years. Moreover, HOS-based identification/equalization methods based on high-order cumulants are theoretically insensitive to Gaussian noise. Many algorithms were studied for the identification/deconvolution of finite-impulse response (FIR) models based on third- or fourth-order cumulants. Nonlinear methods are characterized by equations, where output cumulants and parameters of interests are nonlinearly related. Linear methods use particular cumulant slices, so that the relation becomes linear. It is well known that nonlinear methods can obtain optimum performance (minimum variance), but since the solution of the nonlinear system involves always an iterative search, one is always trapped with local minima of the squared error. Linear methods can obtain good performance when overdetermined systems of equations are used. Most of the works describing HOS-based algorithms, however, were never applied to realistic environments so that eventual advantages of these ideas in the solution of practical problems are not clear. It is well known, just to name the main objection to the use of HOS, that high-order cumulants estimates require considerably more samples than traditional correlation estimates, and, in addition, cumulants estimates are affected by a larger variance of the estimation error. There is an effort in our work in trying to apply an HOS algorithm to a more realistic scenario and to understand the feasibility of practical implementations. The only blind algorithm well studied in its practical implementation [14] is the constant modulus algorithm (CMA) whose demonstrated misconvergence under particular situations constitutes the most important objection.

Manuscript received July 24, 1995; revised November 18, 1996.

The author is with the Watkins-Johnson Company Telecommunications Group, Gaithersburg, MD 20878-1794 USA.

Publisher Item Identifier S 0018-9545(98)02478-5.

The method proposed in this work is based on the same idea introduced in [16], subsequently developed in several algorithms for scalar deconvolution. It is our contribution in this work to generalize some of the mathematical expressions derived in [3] and [20] as they apply to the multichannel, complex case. The resulting algorithm is blind in the sense that it does not require any training sequence or signal to achieve signal detection. Moreover, no angle of arrival needs to be estimated, which makes the method very attractive because the array is not sensitive to calibration errors. The performance of the method is investigated using a square-root approach for the adaptive implementation in the case of a TDMA system with frequency-selective Rayleigh fading. It is important to point out that we focus in this work on the space-time processing of signals severely distorted by multipath frequency-selective fading, and we do not analyze the effect of adjacent channel interference. The multiple-sources scenario can be treated similarly generalizing the main idea introduced here.

The paper is organized as follows. In Section II, we describe the system model for the propagation channel and the discrete-time model. In Section III, the set of equations necessary to solve the deconvolution problem is derived. In Section IV, the adaptive implementation is described, while in Section V the results of the simulations are shown.

II. System Model

We assume a mobile transmitter communicating with a base station with a K-element antenna. Each element of the antenna has a digital filter with $L = L_2 - L_1 + 1$ complex weights. The transfer function of the antenna, whose structure is assumed to be linear with evenly spaced elements, is expressed as

$$H(\omega, \theta) = \sum_{m=L_1}^{L_2} e^{-jm\omega T_d} \sum_{l=1}^{K} w_l(m)\, e^{-jl\phi} \tag{1}$$

where $\phi = (2\pi d \sin\theta/\tilde{\lambda})$, d is the distance between adjacent antenna elements, $\tilde{\lambda}$ is the wavelength of the signal, θ is the arrival angle of the signal (DOA), T_d is the delay between $w_l(m)$ and $w_l(m+1)$ for $l = 1, 2, \cdots, K$ and $m = L_1, L_1 + 1, \cdots, L_2 - 1$, and $H(\omega, \theta)$ is the antenna pattern as a function of θ and the frequency response as a function of ω. Multipath propagation can be characterized as an N-path channel whose ith path $(i = 1, 2, \cdots, N)$ is represented by a P_i received, delayed, and attenuated replica of the signal. The impulse response of the ith path can be expressed as

$$f_i(t) = \sum_{m=1}^{P_i} \rho_{i,m}\, e^{j\psi_{i,m}}\, \delta(t - \tau_{i,m})$$

where $\tau_{i,m}$, $\rho_{i,m}$, and $\psi_{i,m}$ are the delay, amplitude, and phase of the mth delayed signal in the ith path, while $\delta(t)$ is the delta function. Observe that we are assuming now for the derivation a time-invariant channel while instead we should assume that $\tau_{i,m}(t)$, $\rho_{i,m}(t)$, and $\psi_{i,m}(t)$ are time-varying parameters. 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 IV is designed for time-variant channels. The complex baseband modulated signal is $m(t) = \Sigma_k\, x(k)\, p_{tx}(t - kT)$, where $x(k) = a_k + jb_k$ are the complex symbols defining the signal constellation used for the particular digital modulation scheme. $p_{tx}(t)$ is a square-root-raised cosine-shaping filter with rolloff factor 0.35, and T is the signaling interval. The received signal through the ith path can be represented as

$$\begin{aligned} U_i(t) &= \int_{-\infty}^{+\infty} f_i(t-\tau) m(\tau)\, d\tau \\ &= \sum_{m=1}^{P_i} \rho_{i,m} e^{j\psi_{i,m}}\, m(t - \tau_{i,m})\, e^{j2\pi f_0 (t - \tau_{i,m})} \end{aligned} \tag{2}$$

where $\omega_0 = 2\pi f_0$ is the carrier frequency. The contribution of the signal propagated through the ith path with angle of arrival θ_i and phase difference $e^{-j2\pi k d \frac{\sin\theta_i}{\tilde{\lambda}}}$ from the first antenna element to the lth element can be written (we are neglecting the additive noise term) as

$$\begin{aligned} r_l(t, \theta_i) = e^{j\omega_0 t} &\sum_{m=1}^{P_i} \rho_{i,m} m(t - \tau_{i,m}) \\ &\cdot e^{-j2\pi l d (\sin\theta_i/\tilde{\lambda})}\, e^{j\phi_{i,m}} \end{aligned}$$

where $\phi_{i,m} = -2\pi f_0 \tau_{i,m} + \psi_{i,m}$. Sampling at symbol rate T, we can compact the effect of the RF propagation channels at the input of the digital filters at baseband as

$$\begin{aligned} y_i(n) = \sum_m &\, h_i(m) x(n-m) + \eta_i(n) \\ & i = 1, 2, \cdots, K \end{aligned} \tag{3}$$

where $\eta_i(n)$ is Gaussian noise and $h_i(m) = h_i(mT)$ is the T-sampled impulse response

$$\begin{aligned} h_i(t) = \sum_{l=1}^{N} \sum_{m=1}^{P_l} &\rho_{l,m} r_p(t - \tau_{l,m})\, e^{j\phi_{l,m}} \\ &\cdot e^{-j2\pi i d (\sin\theta_l/\tilde{\lambda})}. \end{aligned}$$

In this expression, $r_p(t)$ is the raised cosine function with an excess bandwidth of 0.35 [2] obtained because we assume that the receiver filters $p_{rx}(t)$ at each antenna are square-root-raised cosine filters perfectly matched to the transmitter filters $p_{tx}(t)$. In the z domain, the transfer function (3) can be expressed as

$$\tilde{\mathcal{H}}(z) = \sum_{\nu=-\infty}^{\infty} \boldsymbol{H}(\nu) z^{-\nu} \tag{4}$$

where the organization of the $\mathcal{H}_i(z)$ polynomials (z transforms of $h_i(k)$) in $\tilde{\mathcal{H}}(z)$ is given by

$$\tilde{\mathcal{H}}(z) = [\mathcal{H}_1(z) \quad \mathcal{H}_2(z) \quad \cdots \quad \mathcal{H}_K(z)]^T.$$

If we stack the outputs of the sensors according to the organization of $\mathcal{H}(z), \boldsymbol{y}(t) = (y_1(t), \cdots, y_K(t))^T$, whose z-transform vector polynomial is $\mathcal{Y}(z)$, and we neglect the additive noise term, we can write

$$\mathcal{Y}(z) = \tilde{\mathcal{H}}(z)\mathcal{X}(z)$$

where $\mathcal{X}(z)$ is the z transform of $x(n)$. This is a one-input K-output multichannel moving average model (or SIMO).

A. Distortionless Reception

To recover the input signals, a linear K-input one-output filter as in (1) with tap spacing $T_d = T$ and defined as $\tilde{\mathcal{W}}(z) = \Sigma_{\nu=L_1}^{L_2} \boldsymbol{W}(\nu)z^{-\nu}$ is applied to the downconverted and filtered outputs of the sensors. The main objective for $\tilde{\mathcal{W}}(z)$ is to achieve *distortionless reception*. If we define

$$\tilde{\mathcal{W}}(z) = [\mathcal{W}_1(z) \quad \mathcal{W}_2(z) \quad \cdots \quad \mathcal{W}_K(z)]$$

distortionless reception means that

$$\tilde{\mathcal{W}}(z)\tilde{\mathcal{H}}(z) = 1. \tag{5}$$

The system $\tilde{\mathcal{W}}(z)$ is required to be bounded-input–bounded-output (BIBO) stable. The solution (5) is achievable only ideally. Since the input signal constellations are symmetric, the statistics of the input signal $x(t)$ 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 and a delay. The best possible result for practical *distortionless reception* by a means of a linear filter is

$$\tilde{\mathcal{W}}(z)\tilde{\mathcal{H}}(z) = \mathcal{D}(z) \tag{6}$$

where

$$\mathcal{D}(z) = e^{j\phi_0}z^{-n_0} \ \phi_0 \in [-\pi, \pi], n_0 \text{ integer}.$$

We say that $\tilde{\mathcal{H}}(z)$ satisfies the *distortionless reception* condition if for $\tilde{\mathcal{H}}(z)$ there exists a BIBO stable *distortionless reception* filter $\tilde{\mathcal{W}}(z)$.

III. Description of the Method

We shortly recall that the high-order statistical properties of a process are commonly described in the time domain by cumulants [6]. Cumulants of interest here are fourth-order cumulants of complex zero-mean stationary processes $x(n)$ defined as

$$\begin{aligned}
&\text{cum}\,[x(n_1), x(n_2), x^*(n_3), x^*(n_4)]\\
&\quad = E\{x(n_1)x(n_2)x(n_3)^*x(n_4)^*\}\\
&\qquad - E\{x(n_1)x(n_2)\}\ E\{x(n_3)^*x(n_4)^*\}\\
&\qquad - E\{x(n_1)x(n_3)^*\}\ E\{x(n_2)x(n_4)^*\}\\
&\qquad - E\{x(n_1)x(n_4)^*\}\ E\{x(n_2)x(n_3)^*\}.
\end{aligned}$$

Second-order cumulants are defined as

$$\text{cum}\,[x(n_1), x^*(n_2)] = E\{x(n_1)x(n_2)^*\}.$$

The properties of cumulants that we exploit are:

1) *LIN*

$$\text{cum}\,[\Sigma_n\ f(n)x(n), \cdots] = \Sigma_n\ f(n)\ \text{cum}\,[x(n), \cdots,];$$

2) *STATIND* if the samples of a process can be divided into two (or more) statistically independent subsets and 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 $k > 2$.

A. Key Assumptions

The important assumptions necessary to derive the algorithm are as follows.

1) AS1: the transformation in (3) represents a stable system, but possibly NMP, satisfying:
   a) all channels $h_i(k), i = 1, 2, \cdots, K$, which have finite support q;
   b) $h_i(0) \neq 0$ for some i;
   c) $h_i(q) \neq 0$ for some i;
   d) $\mathcal{H}_i(z) = \Sigma_{\nu=0}^{q}\ h_i(\nu)\ z^{-\nu}$, for $i = 1, 2, \cdots, K$, which have no common zeroes.
2) AS2: the complex sequence $\{x(n)\}$ is constituted by random variables identically non Gaussian distributed and statistically independent, and the cumulants of $\{x(n)\}$ satisfy:
   a) cum $[x(n), x^*(n)] = \gamma_2 > 0$;
   b) cum $[x(n), x^*(n), x(n), x^*(n)] = \gamma_4 \neq 0$;
   c) cum $[x(n), x(n)] = $ cum $[x^*(n), x^*(n)] = 0$.

Assumption AS1 is necessary because it assures the *distortionless reception* condition for $\tilde{\mathcal{H}}(z)$. The relation between $\tilde{\mathcal{H}}(z)$ and $\tilde{\mathcal{W}}(z)$ can also be expressed as

$$\boldsymbol{\xi} = H_{L,q}\overline{\boldsymbol{w}} \tag{7}$$

where

$$\boldsymbol{\xi} = (\cdots, \xi(-1), \xi(0), \xi(1), \cdots)^T$$

is the $1 \times (L+q)$ impulse response of the cascaded system $\tilde{\mathcal{W}}(z)\tilde{\mathcal{H}}(z)$

$$\overline{\boldsymbol{w}} = (\boldsymbol{w}_1^T, \boldsymbol{w}_2^T, \cdots, \boldsymbol{w}_K^T)^T, \qquad \boldsymbol{w}_i = (w_i(L_1), \cdots, w_i(L_2))^T$$

and $1 \times KL.H_{L,q}$ is the $(L+q) \times KL$ generalized Sylvester matrix

$$H_{L,q} = [H_{L,q}^{(1)} \quad H_{L,q}^{(2)} \quad \cdots \quad H_{L,q}^{(K)}]$$

$$H_{L,q}^{(i)} = \begin{bmatrix} h_i(0) & h_i(1) & \cdots & h_i(q) & 0 & \cdots \\ 0 & h_i(0) & h_i(1) & \cdots & h_i(q) & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & 0 & h_i(0) & h_i(1) & \cdots & h_i(q) \end{bmatrix}^T.$$

The rank of $H_{L,q}$ plays a crucial role in the existence of a stable BIBO *distortionless reception* filter $\tilde{\mathcal{W}}(z)$. The relationship between the rank of the generalized Sylvester matrix $H_{L,q}$ and the *reducibility* of $\tilde{\mathcal{H}}(z)$ has been studied in multivariable control literature [9], [10]. An $m \times n$ polynomial matrix $\mathcal{M}(z)$ (with $m \geq n$) is said to be *irreducible* if there is no $n \times n$ matrix $\mathcal{B}(z)$ with a nonconstant determinant such that $\mathcal{M}(z) = \tilde{\mathcal{M}}(z)\mathcal{B}(z)$, where $\tilde{\mathcal{M}}(z)$ is also $m \times n$. It is proved in [9] that $\tilde{\mathcal{H}}(z)$ is irreducible if and only if it is full rank for any z (which is expressed in our case by AS1).

The desired solution $\boldsymbol{\xi}$ that completely restores the information signal up to the delay n_0 is $\boldsymbol{\delta}$. The generic mth element of the vector $\boldsymbol{\delta}$ is $\delta(m)$, if we neglect the phase shift ϕ_0 and the delay n_0. Formally, we have to solve the minimization problem

$$\min_{\overline{\boldsymbol{w}}} \ ||H_{L,q}\overline{\boldsymbol{w}} - \boldsymbol{\delta}||^2. \tag{8}$$

This is a linear least-squares estimation problem, whose solution requires $H_{L,q}$ to be full rank. First, we observe that by AS1, $\tilde{\mathcal{H}}(z)$ is irreducible. Second, we specify the following lemma.

Lemma 1: If AS1 is valid, $H_{L,q}$ is full rank for any $L = L_2 - L_1 + 1 \geq \lceil qK/K - 1 \rceil$.

Proof: Let us assume that $\boldsymbol{H}(q)$ is constituted by $b_l(q)$, $l = 1, 2, \cdots, K$ elements and that $b_i(q)$ is an element of $\boldsymbol{H}(q)$ different from zero (which exists due to AS1). Let $\tilde{b}(n) = (b_i(n)$ for $n = 0 \cdots, q-1$ and $\mathcal{B}(z) = \Sigma_{n=0}^{q} \tilde{b}(n) z^{-n}$. The degree of $\mathcal{B}(z)$ is q. By Corollary 1 of [10], $\mathcal{H}(z)$ is irreducible if and only if $L \geq \lceil qK/K - 1 \rceil$ and

$$\text{rank } \{H_{L,q}\} = L + \text{degree}(\mathcal{B}(z)) = L + q$$

which means that $H_{L,q}$ is full rank for $L = L_2 - L_1 + 1 \geq \lceil qK/K - 1 \rceil$. □

The meaning of this result is that if we choose the appropriate length L, $H_{L,q}^{T^*} H_{L,q}$ is invertible, and for *any* $\boldsymbol{\xi}$, there exists

$$\overline{\boldsymbol{w}} = (H_{L,q}^{T^*} H_{L,q})^{-1} H_{L,q}^{T^*} \boldsymbol{\xi}$$

such that $\boldsymbol{\xi} = H_{L,q} \overline{\boldsymbol{w}}$, which is equivalent to say that the solution of (8) exists as

$$\overline{\boldsymbol{w}}^{\text{opt}} = (H_{L,q}^{T^*} H_{L,q})^{-1} H_{L,q}^{T^*} \boldsymbol{\delta}.$$

In the time domain, the filter can be expressed as

$$z(n) = \sum_{i=1}^{K} \sum_{m=L_1}^{L_2} w_i(m) y_i(n-m) \tag{9}$$

and for $\overline{\boldsymbol{w}} = \overline{\boldsymbol{w}}^{\text{opt}}$, $\boldsymbol{\xi} = H_{L,q} \overline{\boldsymbol{w}}^{\text{opt}} \simeq \boldsymbol{\delta}$, hence

$$z(n) = \sum_{i=1}^{K} \sum_{m=L_1}^{L_2} w_i^{\text{opt}}(m) y_i(n-m) \simeq x(n). \tag{10}$$

As L is made larger, the matrix $H_{L,q}(H_{L,q}^{T^*} H_{L,q})^{-1} H_{L,q}^{T^*}$ tends to the identity matrix and (10) tends to equality.

B. Derivation

We assume, for the derivation, that there is no noise in the model. Using statistical independence of the source symbols $\{x(n)\}$ (STATIND) and the linearity property of cumulants (LIN), we can write for the following fourth-order cross moments:

$$\begin{aligned}
E\{x(k) & y_{i_1}(k+n_1)^* y_{i_2}(k+n_1) y_{i_3}(k+n_1)^*\} \\
&= \sum_{m=0}^{q} \sum_{n=0}^{q} \sum_{n_2=0}^{q} h_{i_1}(m)^* h_{i_2}(n) h_{i_3}(n_2)^* \\
&\quad \times E\{x(k) x(k+n_1-m)^* x(k+n_1-n) \\
&\quad \cdot x(k+n_1-n_2)^*\} \\
&= \gamma_4 h_{i_1}(n_1)^* h_{i_2}(n_1) h_{i_3}(n_1)^* \\
&\quad + \gamma_2^2 \left[\sum_{m=0}^{q} h_{i_1}(m)^* h_{i_2}(m) \right] h_{i_3}(n_1)^* \\
&\quad + \gamma_2^2 \left[\sum_{n=0}^{q} h_{i_b}(n) h_{i_c}(n)^* \right] h_{i_a}(n_1)^*.
\end{aligned} \tag{11}$$

This expression is evidently different from zero only for $0 \leq n_1 \leq q$. According to the assumed model and since we are looking for the noncausal filter that verifies (10), we can write

$$\begin{aligned}
E\{x(k) & y_{i_2}(k+n_1)^* y_{i_3}(k+n_1) y_{i_4}(k+n_1)^*\} \\
&\simeq E\{z(k) y_{i_2}(k+n_1)^* y_{i_3}(k+n_1) y_{i_4}(k+n_1)^*\} \\
&= \sum_{i_1=1}^{K} \sum_{k_1=L_1}^{L_2} w_{i_1}(k_1) \\
&\quad \times E\{y_{i_1}(k-k_1) y_{i_2}(k+n_1)^* \\
&\quad y_{i_3}(k+n_1) y_{i_4}(k+n_1)^*\}.
\end{aligned} \tag{12}$$

The fourth-order moments are related to the fourth- and second-order cumulants [20] (in our case, the odd moments are assumed to be zero) as

$$\begin{aligned}
E\{y_{i_1} & (k-k_1) y_{i_2}(k+n_1)^* y_{i_3}(k+n_1) y_{i_4}(k+n_1)^*\} \\
&= \text{cum } [y_{i_1}(k), y_{i_2}(k+n_1+k_1)^*, y_{i_3}(k+n_1+k_1) \\
&\quad y_{i_4}(k+n_1+k_1)^*] \\
&\quad + \text{cum } [y_{i_1}(k), y_{i_2}(k+n_1+k_1)^*] \\
&\quad \cdot \text{cum}[y_{i_3}(k), y_{i_4}(k)^*] \\
&\quad + \text{cum } [y_{i_1}(k), y_{i_4}(k+n_1+k_1)^*] \\
&\quad \cdot \text{cum } [y_{i_2}(k)^*, y_{i_3}(k)].
\end{aligned} \tag{13}$$

Using cum $[y_{i_1}(k)^*, y_{i_2}(k)] = \gamma_2 \sum_{n=0}^{q} h_{i_1}(n)^* h_{i_2}(n)$ into (13), then substituting (13) into (14) and using

$$\begin{aligned}
\sum_{i_1=1}^{K} & \sum_{k_1=L_1}^{L_2} w_{i_1}(k_1) c_{i_1,i_2}(n_1+k_1) \\
&= \begin{cases} \gamma_2 h_{i_2}(n_1)^*, & n_1 \in [0, q] \\ 0, & \text{otherwise} \end{cases}
\end{aligned} \tag{14}$$

where $c_{i_1,i_2}(n_1+k_1) = \text{cum } [y_{i_1}(k), y_{i_2}(k+n_1+k_1)^*]$ [see Appendix B for the derivation of (14)], we obtain

$$\begin{aligned}
E\{x(k) & y_{i_2}(k+n_1)^* y_{i_3}(k+n_1) y_{i_4}(k+n_1)^*\} \\
&= \sum_{i_1=1}^{K} \sum_{k_1=L_1}^{L_2} w_{i_1}(k_1) c_{i_1,i_2,i_3,i_4}(k_1+n_1) \\
&\quad + \gamma_2^2 \left[\sum_{n=0}^{q} h_{i_2}(n)^* h_{i_3}(n) \right] h_{i_4}(n_1) \\
&\quad + \gamma_2^2 \left[\sum_{n=0}^{q} h_{i_3}(n) h_{i_4}(n)^* \right] h_{i_2}(n_1)^*
\end{aligned} \tag{15}$$

if $n_1 \in [0, q]$ and

$$\begin{aligned}
E\{x(k) & y_{i_2}(k+n_1)^* y_{i_3}(k+n_1) y_{i_4}(k+n_1)^*\} \\
&= \sum_{i_1=1}^{K} \sum_{k_1=L_1}^{L_2} w_{i_1}(k_1) c_{i_1,i_2,i_3,i_4}^{y}(k_1+n_1)
\end{aligned} \tag{16}$$

in any other case with

$$\begin{aligned}
c_{i_1,i_2,i_3,i_4}^{y} & (m) \\
&= \text{cum } [y_{i_1}(k), y_{i_2}(k+m)^*, y_{i_3}(k+m) \\
&\quad y_{i_4}(k+m)^*]
\end{aligned}$$

$$c_{i_1,i_2}^{y}(m) = \text{cum } [y_{i_1}(k), y_{i_2}(k+m)^*].$$

Equating (15)–(16) and (11), we obtain

$$\sum_{i_1=1}^{K}\sum_{k_1=L_1}^{L_2} w_{i_1}(k_1)c^y_{i_1,i_2,i_3,i_4}(k_1+n_1) = \begin{cases}\gamma_4 h_{i_2}(n_1)^* h_{i_3}(n_1) h_{i_4}(n_1)^*, & \text{if } 0 \le n_1 \le q\\ 0, & \text{otherwise.}\end{cases} \tag{17}$$

Expression (17) is important because it relates linearly the weights $w_i(k)$ to the fourth-order cross cumulants of the samples of the signal (filtered and downconverted) at the output of the antenna elements. However, we cannot use (17) directly because the terms $h_{i_2}(n_1)^* h_{i_3}(n_1) h_{i_4}(n_1)^*$ in the linear equations require knowledge of the baseband discrete-time channels not available in practical systems. To eliminate the terms containing the unknown channels, we can consider a derivation similar to the Giannakis–Mendel (GM) method [5]. The original GM method was designed for the real-time-series case, so we have to make some modifications to apply the idea to the complex multichannel case. Following the derivation of Appendix C, we finally obtain

$$\begin{aligned}\sum_{m=0}^{q}\sum_{i_1=1}^{K}\sum_{k_1=L_1}^{L_2} & w_{i_1}(k_1)c^y_{i_1,l}(m+k_1)c^y_{i,l,l,l}(n-m)\\ &= \sum_{m=0}^{q}\sum_{i_1=1}^{K}\sum_{k_1=L_1}^{L_2} w_{i_1}(k_1)c^y_{i_1,l,l,l}(k_1+m)c^y_{i,l}(n-m)\\ &\quad n=-q,\cdots,2q, \qquad i,l=1,2,\cdots,K.\end{aligned} \tag{18}$$

This is an overdetermined system of equations with $(L_2-L_1+1)K$ unknowns and $(3q+1)K^2$ equations. Assuming $w_1(0)=1$, we can collect these equations in a system, linear in the coefficients $w_i(k_1)$ $(i_1 \neq 1, k_1 \neq 0)$. We define the following statistics:

$$\begin{aligned}s^{i,l}_{i_1,k_1,n} &= \sum_{m=0}^{q}[c^y_{i_1,l}(k_1+m)c^y_{i,l,l,l}(n-m)\\ &\quad - c^y_{i,l}(n-m)c^y_{i_1,l,l,l}(k_1+m)]\\ &\quad n=-q,-q+1,\cdots,2q, i,l,i_1=1,2,\cdots,K\\ &\quad k_1=L_1,\cdots,L_2.\end{aligned}$$

Compacting the unknown weights in vector notation

$$\begin{aligned}\boldsymbol{w} &= [\tilde{\boldsymbol{w}}^T, \boldsymbol{w}_2^T, \cdots \boldsymbol{w}_K^T]^T\\ \boldsymbol{w}_i &= [w_i(L_1), w_i(L_1+1), \cdots w_i(L_2)]^T\\ \tilde{\boldsymbol{w}} &= [w_1(L_1), w_1(L_1+1), \cdots w_1(-1), w_1(1), \cdots w_1(L_2)]^T\end{aligned}$$

the system of equations can be written

$$\boldsymbol{C}\boldsymbol{w} = \boldsymbol{d} \tag{19}$$

where the matrix $\boldsymbol{C}$ is defined as

$$\boldsymbol{C} = [\boldsymbol{C}^T_{1,1} \quad \boldsymbol{C}^T_{1,2} \quad \cdots \quad \boldsymbol{C}^T_{1,K} \quad \boldsymbol{C}^T_{2,1} \quad \cdots \quad \boldsymbol{C}^T_{K,K}]^T$$

and

$$\boldsymbol{C}^T_{i,l} = [\tilde{\boldsymbol{S}}^{i,l} \quad \boldsymbol{S}^{i,l}_2 \quad \cdots \quad \boldsymbol{S}^{i,l}_K]$$

with the equation, given at the bottom of the page, and

$$\boldsymbol{S}^{i,l}_n = \begin{bmatrix} s^{i,l}_{n,L_1,-q} & s^{i,l}_{n,L_1+1,-q} & \cdots & s^{i,l}_{n,L_2,-q}\\ s^{i,l}_{n,L_1,-q+1} & s^{i,l}_{n,L_1+1,-q+1} & \cdots & s^{i,l}_{n,L_2,-q+1}\\ \vdots & \vdots & \vdots & \vdots\\ s^{i,l}_{n,L_1,2q} & s^{i,l}_{n,L_1+1,2q} & \cdots & s^{i,l}_{n,L_2,2q}\end{bmatrix}.$$

The vector $\boldsymbol{d}$ is defined as

$$\boldsymbol{d} = [\boldsymbol{d}^T_{1,1} \quad \boldsymbol{d}^T_{1,2} \quad \cdots \quad \boldsymbol{d}^T_{1,K} \quad \boldsymbol{d}^T_{2,1} \quad \cdots \quad \cdots \quad \boldsymbol{d}^T_{K,K}]^T$$

with

$$\boldsymbol{d}_{i,l} = [-s^{i,l}_{1,0,-q} \quad -s^{i,l}_{1,0,-q+1} \quad \cdots \quad -s^{i,l}_{1,0,2q}]^T.$$

This system can be solved in the least-squares sense

$$\boldsymbol{w} = \boldsymbol{C}^\dagger \boldsymbol{d} = (\boldsymbol{C}^H\boldsymbol{C})^{-1}\boldsymbol{C}^H\boldsymbol{d}$$

because of the assumptions. Particularly, $AS1$ assures that the *distortionless reception* condition is satisfied so that the model (10) is valid, and, as a consequence, $\boldsymbol{C}^H\boldsymbol{C}$ is positive definite, [3], [5]. When the samples are well separated in time and if the cumulants are absolutely summable, then the theoretical cumulants are consistently estimated from a data record of N samples, and ensemble averages can be approximated by empirical averages.

IV. Adaptive Implementation Using an RLS Square-Root Approach

We can consider a recursive implementation if we adopt the cumulant estimation procedure reported in Appendix A. Defining as $\boldsymbol{C}(n), \boldsymbol{w}(n), \boldsymbol{d}(n)$ the estimation of $\boldsymbol{C}, \boldsymbol{w}, \boldsymbol{d}$, respectively, at time instant n, we can devise the following adaptive algorithm. We denote the ith row of a matrix $\boldsymbol{M}$ as $[\boldsymbol{M}]_i$. Similarly, for vectors we denote the ith element of a vector $\boldsymbol{v}$ as $[\boldsymbol{v}]_i$. The matrix $\tilde{\boldsymbol{C}}(n)$ and the vector $\tilde{\boldsymbol{d}}(n)$ are defined by the following time updates:

$$\tilde{\boldsymbol{C}}(n+1) = \begin{pmatrix}\lambda\tilde{\boldsymbol{C}}(n)\\ [\boldsymbol{C}(t[n+1])]_{i[n+1]}\end{pmatrix}$$

$$\tilde{\boldsymbol{d}}(n+1) = \begin{pmatrix}\lambda\tilde{\boldsymbol{d}}(n)\\ [\boldsymbol{d}(t[n+1])]_{i[n+1]}\end{pmatrix}$$

$$\tilde{\boldsymbol{S}}^{i,l} = \begin{bmatrix} s^{i,l}_{1,L_1,-q} & \vdots & s^{i,l}_{1,-1,-q} & s^{i,l}_{1,1,-q} & \vdots & s^{i,l}_{1,L_2,-q}\\ s^{i,l}_{1,L_1,-q+1} & \vdots & s^{i,l}_{1,-1,-q+1} & s^{i,l}_{1,1,-q+1} & \vdots & s^{i,l}_{1,L_2,-q+1}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ s^{i,l}_{1,L_1,2q} & \vdots & s^{i,l}_{1,-1,2q} & s^{i,l}_{1,1,2q} & \vdots & s^{i,l}_{1,L_2,2q}\end{bmatrix}$$

TABLE I
COMPUTATIONAL COMPLEXITY FOR K ELEMENTS, L TAPS, AND CHANNEL IMPULSE RESPONSES TRUNCATED TO $q+1$ FOR THE HOS ALGORITHM AND THE QR-RLS

Method	HOS	QR-RLS
N. of real multiplies	$8K^3L(3q+1)(q+1)+36K^2(3q+1)+10(KL)^2+30KL$	$10(KL)^2+30KL$
N. of reciprocals	$2KL$	$2KL$
N. of Square Roots	KL	KL

where $t[n] = \lfloor n - 1/K^2(3q+1)\rfloor + 1$, $i[n] = (n-1)\mathrm{mod}[K^2(3q+1)] + 1$, and λ is the forgetting factor. At each $n+1$ stage, we wish to solve the problem

$$\min_{\boldsymbol{w}} \left\| \begin{pmatrix} \lambda\tilde{\boldsymbol{C}}(n) \\ [\boldsymbol{C}(t[n+1])]_{i[n+1]} \end{pmatrix} \boldsymbol{w} - \begin{pmatrix} \lambda\tilde{\boldsymbol{d}}(n) \\ [\boldsymbol{d}(t[n+1])]_{i[n+1]} \end{pmatrix} \right\|^2. \tag{20}$$

Suppose that a matrix $\boldsymbol{V}(n)$ is known, such that $\boldsymbol{Q}^H\tilde{\boldsymbol{C}}(n) = \binom{\boldsymbol{V}(n)}{0}$ with $\boldsymbol{Q}$ orthogonal and $\boldsymbol{V}(n)$ upper triangular, then the problem stated in (20) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{w}} & \left\| \boldsymbol{Q}^H \begin{pmatrix} \lambda\boldsymbol{C}(n) \\ [\boldsymbol{C}(t[n+1])]_{i[n+1]} \end{pmatrix} \boldsymbol{w} - \boldsymbol{Q}^H \begin{pmatrix} \lambda\tilde{\boldsymbol{d}}(n) \\ [\boldsymbol{d}(t[n+1])]_{i[n+1]} \end{pmatrix} \right\|^2 \\ = \min_{\boldsymbol{w}} & \left\| \begin{pmatrix} \lambda\boldsymbol{V}(n) \\ [\boldsymbol{C}(t[n+1])]_{i[n+1]} \end{pmatrix} \boldsymbol{w} - \begin{pmatrix} \lambda\bar{\tilde{\boldsymbol{d}}}(n) \\ [\boldsymbol{d}(t[n+1])]_{i[n+1]} \end{pmatrix} \right\|^2 \end{aligned} \tag{21}$$

because the Euclidean distance is preserved by orthogonal transformations. The proof of (21) can be derived forming the *normal equations* for the two minimization problems and comparing them. The advantage is that the solution minimizer of (21) is simply the solution of a triangular system. We update for the change in the parameter (see [17]) $d\boldsymbol{w}(n) = \boldsymbol{w}(n+1) - \boldsymbol{w}(n)$. To find the orthogonal matrix $\boldsymbol{Q}$, an efficient procedure can be adopted similar to the technique proposed in [17]: a set of Givens rotations to annihilate the inferior triangular part of the matrix $\binom{\lambda\tilde{\boldsymbol{C}}(n)}{[\boldsymbol{C}(t[n+1])]_{i[n+1]}}$. The algorithm consists of the following steps.

1) Compute the prediction error $u(n+1) = [\boldsymbol{d}(t[n+1])]_{i[n+1]} - [\boldsymbol{C}(t[n+1])]^T_{i[n+1]}\boldsymbol{w}(n)$.
2) Form the matrix
$$\begin{pmatrix} \lambda\boldsymbol{V}(n) & \mathbf{o} \\ [\boldsymbol{C}(t[n+1])]_{i[n+1]} & u(n+1) \end{pmatrix}.$$
3) Sweep the bottom part of this matrix using Givens rotations.
4) Solve the triangular system $\boldsymbol{V}(n+1)d\boldsymbol{w}(n+1) = \bar{\tilde{\boldsymbol{d}}}(n+1)$.

A. Remarks on the Implementation

The simulations in the following section are performed using a word length size of 24 bits using fixed-point arithmetic. Digital signal processing processors are commercially available with these characteristics. As expected, the square-root (QR)-based filter exhibited excellent numerical behavior and robustness to roundoff errors. We also reduced the number of bits to 16 to use fixed-point arithmetic processors with different costs without observing important problems of convergence. The computational complexity in terms of multiplications per iteration was calculated and compared to the complexity of the adaptive QR-RLS (recursive least squares based on QR decomposition) approach summarized in Appendix D. To make a fair comparison, we used the same square-root approach in the implementation of the RLS algorithm. So, while in the HOS case square-root decompositions are performed on a cumulant matrix, in the case of the MMSE, the square-root decomposition is performed on the data samples properly organized into a matrix (see Appendix D). The approximate number of computations per iteration using filters length equal to L is given in Table I. The Givens rotations can be optimized and modified into "fast" Givens rotations to eliminate the square roots. However, we observe that "fast" Givens rotations, if computationally more efficient, are known to introduce some numerical instability. On the other hand, square roots can be easily implemented in a DSP processor by lookup tables.

TABLE II
EXAMPLE OF COMPUTATIONAL COMPLEXITY FOR $K=4$ AND $L=9$ TAPS AND CHANNEL IMPULSE RESPONSES TRUNCATED TO $Q+1=5$ FOR THE HOS ALGORITHM AND THE QR-RLS

Method	HOS	QR-RLS
N. of real multiplies ($log_{10}N$)	5.5066	4.1474
N. of reciprocals ($log_{10}N$)	1.6812	1.6812
N. of Square Roots ($log_{10}N$)	1.3802	1.3802

From Tables I and II, it is evident that the HOS algorithm here proposed has a significantly higher computational complexity than the traditional QR-RLS scheme. This is the price to be paid for the improved performance in terms of identification capability and *blindness* of the approach. It is extremely important to emphasize that if throughput rate is of concern, one may use efficient very large-scale integration (VLSI) architectures because an increase in complexity does not necessarily result directly in lower computational speed.

V. SIMULATIONS

To show how the method can be used in a digital mobile radio, system we analyze the performance of an ideal TDMA system (similar to the IS-54 standard) for cellular communications. The mobile to base communication is allocated in the 824–849-MHz band. The 30-KHz channel of the transmitter employs QPSK at a data rate of 13 Kbps. A block diagram of the receiver is shown in Fig. 1. Antenna spacing is $\tilde{\lambda}/2$. The

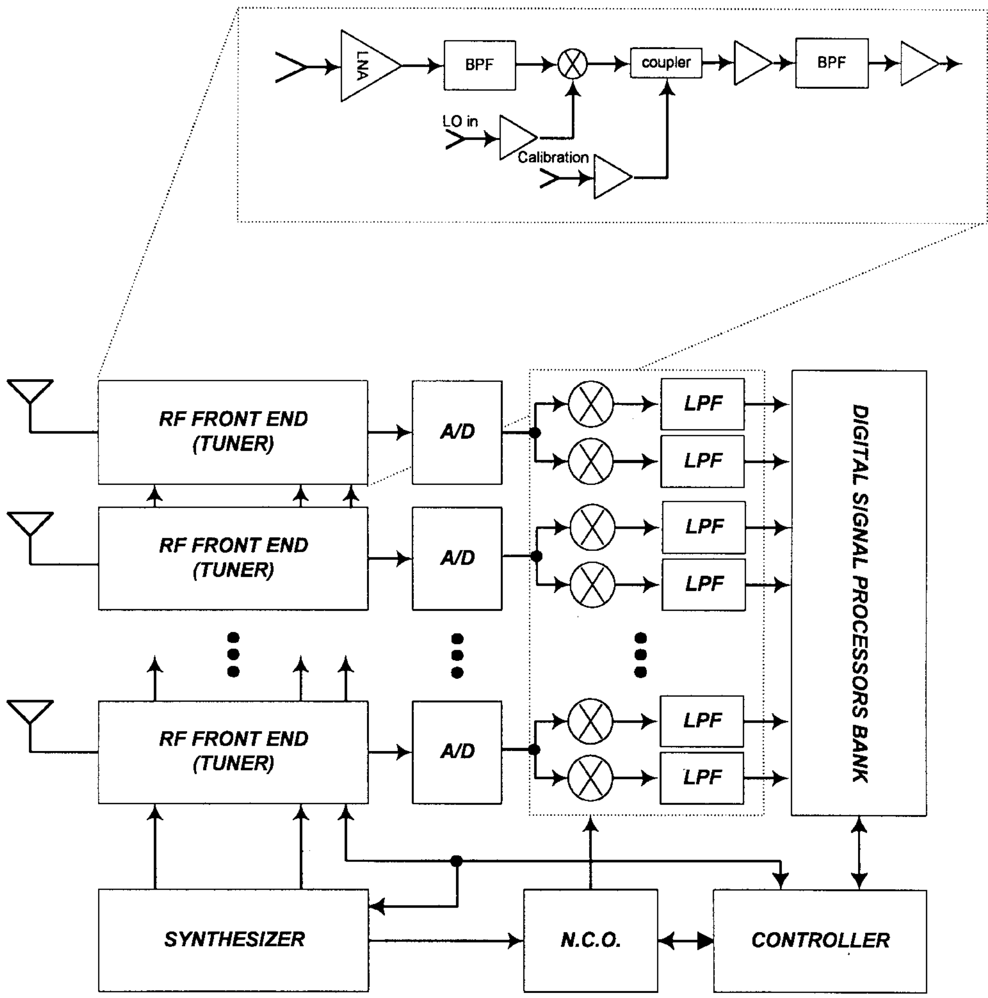


Fig. 1. Block diagram of the receiver.

tuner module performs a standard single conversion scheme. The analog-to-digital (A/D) converter is a high-speed bandpass sampler, while the conversion at baseband is operated by digital downconverters (channelizers). This architecture is cost effective and very flexible, but the performance of the receiver critically depends on the A/D converters. Pulse-shaping filters are square-root-raised cosine filters at the transmitter and at the receiver, so that $r_p(t)$ is a perfect raised cosine function with excess bandwidth equal to 0.35. In Fig. 2, we show the equivalent baseband discrete-time model, where the tapped delay lines are the time-filtering stages. We assume three independent rays are received at the antenna $(N = 3)$, and each ray is characterized by $P_1 = 22$, $P_2 = 20$, and $P_3 = 18$ paths. The powers of the delayed paths, that is, $E\{|\rho_{i,m}|^2\}$ and the delays $\tau_{i,m}$ for $m = 1, 2, \cdots P_i$ and $i = 1, 2, 3$, are distributed according to the power delay profile, which is constituted by two clusters with one-sided exponential delay

$$\Phi_i(\tau) = \begin{cases} e^{-\tau}, & 0 \le \tau \le D_i \\ 0.5e^{(D_i - \tau)}, & D_i \le \tau < 2D_i. \end{cases}$$

where τ and D_i are expressed in microseconds. The values of the actual delays for $i = 1,\ 2$, and 3 are obtained by uniform sampling of $\Phi_i(\tau)$, that is, $\tau_{i,m} = m(2D_i/P_i$, then the power of the rays are $E\{|\rho_{i,m}|^2\} = \Phi_i(\tau_{i,m})$. It is then evident that each channel is specified uniquely by the delay interval D_i, which we normalize with respect to the symbol period T.

The angles of arrival of the three rays are $\theta_1 = -10^\circ, \theta_2 = 10^\circ$, and $\theta_3 = 15^\circ$. The normalized (to symbol period) delay interval of the first path is 0.15, while the second and third paths are specified in each test case. The number of elements of the antenna is four. Tap spacing is equal to one symbol period. The support of each (FIR) channel (the impulse response of each path) is truncated to five, that is, the impulse responses of the symbol-sampled channels span approximately five symbol periods. This means that it is assumed that $q = 4$ in the algorithm. The Doppler frequency usually describes the SOS's of channel variations. Doppler frequency is related through wavelength $\tilde{\lambda}$ to vehicle motion. The model used is based on the wide-sense stationary uncorrelated scattering (WSSUS) assumption. The complex weights $\rho_{i,m}$ are generated as filtered Gaussian processes 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 by 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 channels and complex weights have the same

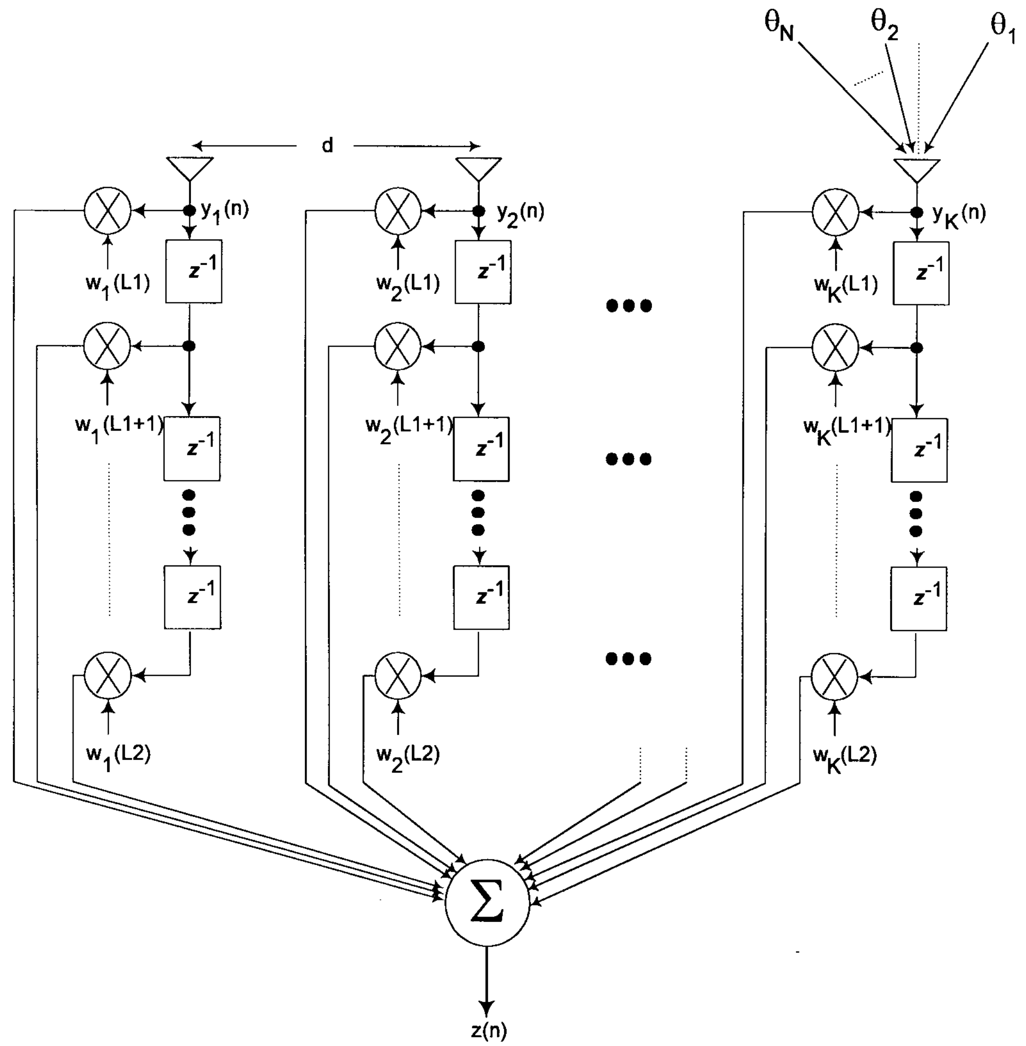


Fig. 2. Discrete-time model of the filtering section (K sensors and N paths).

Doppler spectrum. To show how the use of the traditional MMSE (SOS-based) approach can be inadequate because of the MP requirement, we studied the variations of the equivalent discrete-time model as related to a common propagation condition. Fig. 3 shows the real-time position of the zeroes of one of the four equivalent discrete-time channel models (there are four zeroes because $q = 4$) for a channel model reflecting the moving speed of the mobile of 27 mi/h and a delay interval of 0.5. The dashed circle is the unit circle in the complex plane. It is possible to see that most of the realizations of the channel present a strong NMP condition. The QR-RLS MMSE filter does not perform optimally in these situations, as also shown in the following results, and this can motivate the use of high-order statistics. The channels for Figs. 4 and 5 are static in each run, but they have different realization from run to run according to the statistical distribution of the multipath parameters (the amplitudes $|\rho_{i,m}|$ are Rayleigh distributed, while the phases $\arg(\rho_{i,m})$ are uniformly distributed). The mean-squared error (MSE) is defined as the average of the squared errors obtained over $\tilde{M} = 100$ Monte Carlo runs and is given by $\mathrm{MSE}(n) = (1/\tilde{M})\,\Sigma_{\alpha=1}^{\tilde{M}}\,\overline{e}_\alpha(n)$, where $\overline{e}_\alpha(n) = (1/N_{\mathrm{av}})\,\Sigma_{l=0}^{N_{\mathrm{av}}}\,|e_\alpha(n-l)|^2, e_\alpha(n) = z_\alpha(n) - x(n-n_0), n_0$

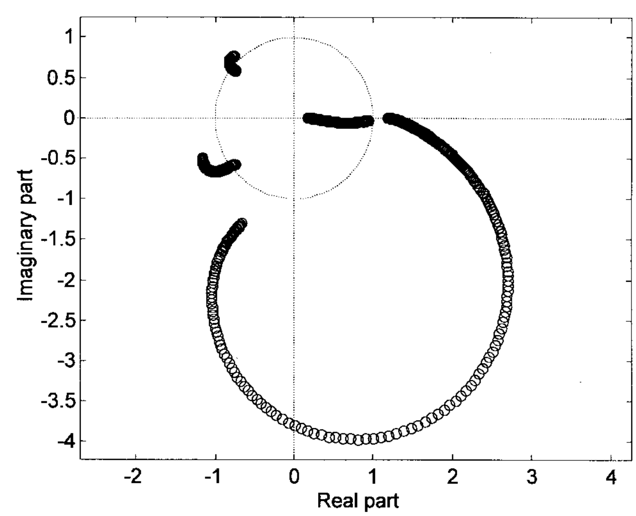


Fig. 3. Trajectory of the zeroes in the complex plane of the equivalent discrete-time model for the Rayleigh fading channel. These plots are representative of a fading event spanning 300 symbols. The speed of the mobile transmitter is 27 mi/h, and the delay interval D_i of the channel is 0.5. The NMP characteristic of the channel during most of the channel realizations is evident.

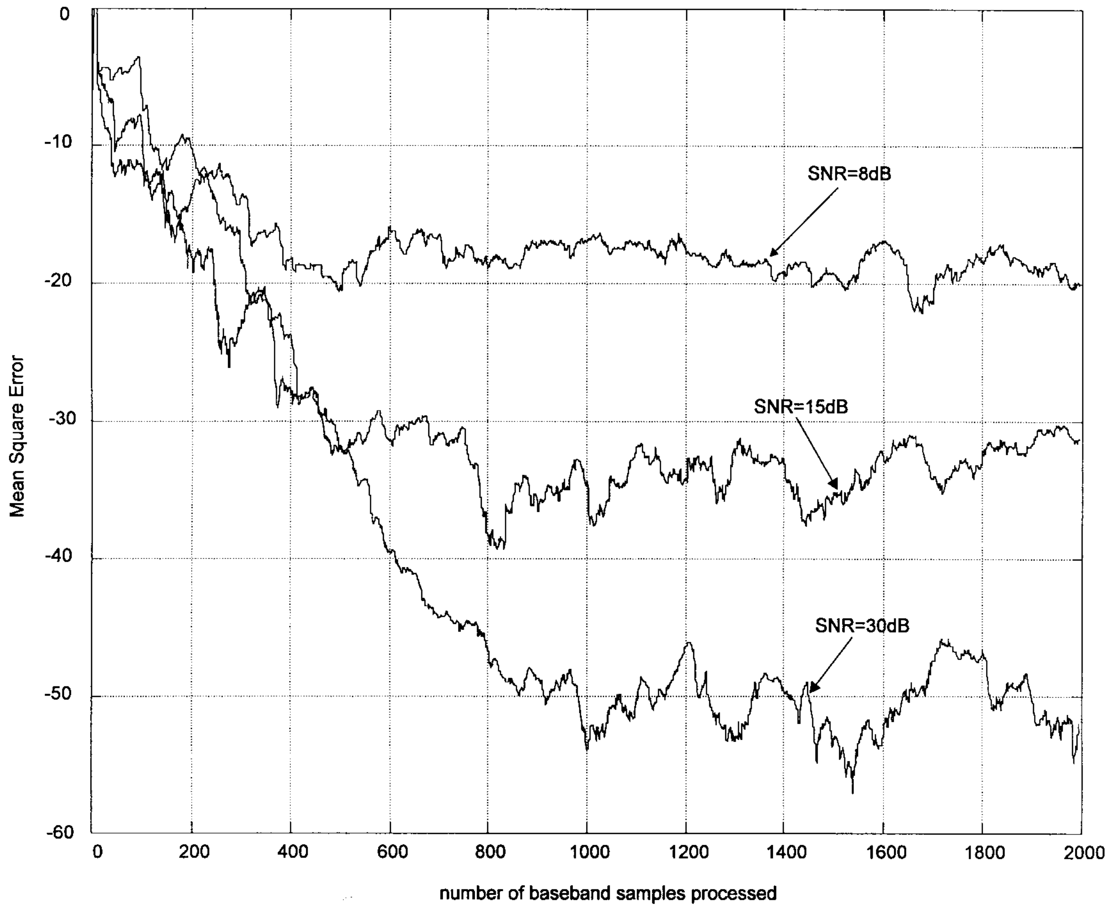


Fig. 4. Convergence process for SNR $= 10$ dB on the second and third paths and different SNR's on the first path. Delay intervals are 0.15, 0.25, and 0.35.

is the delay introduced by the filters, and $z_\alpha(n)$ is the output $z(n)$ of the combined filters obtained at the αth run. $N_{\rm av}$ introduces a short-term average of the MSE. In the simulation results $N_{\rm av} = 10$.

Fig. 4 shows the results of the simulations in terms of the MSE when using $L = 30, L_1 = -15, L_2 = 15$ (the delay intervals $D_2/T$and D_3/T for the second and third paths are 0.25 and 0.35), and $\lambda = 1$. Signal-to-noise ratio (SNR) for each discrete-time channel impulse response is defined as in [1]. The performance of the convergence process of the estimation algorithm is shown at different SNR's. In Fig. 5, for the same delay intervals and $\lambda = 1$ and SNR's environment of the previous test case we increase the length of the filters from 25 to 40. It is evident that the computational complexity increase does not justify the marginal improvement.

To evaluate the benefits obtained using high-order statistics as opposed to the traditional MMSE approach, we compared the performance of the QR-RLS method described in Appendix D and the HOS algorithm. We used $L = 16,\ L_1 = -6$, and $L_2 = 9$ and delay intervals $D_1/T = 0.15,\ D_2/T = 0.25$, and $D_3/T = 0.5$ and $D_1/T = 0.15,\ D_2/T = 0.5$, and $D_3/T = 0.75$, respectively, for the first, second, and third paths, respectively, in Fig. 6(a) and (b). The SNR is 10 dB, and $\lambda = 1$. The results of this experiment are reported in Fig. 6. The HOS algorithm can reach lower values of the MSE, but a definitely slower rate of convergence. The lower MSE is due to the improved identification capability of higher than SOS's, and the slow convergence makes evident the well-known requirement for larger sample size to estimate the HOS.

Fig. 7 shows the impulse response frequency response of the discrete-time model before and after convergence of the weights in the same situation of Fig. 5 and SNR $= 30$ dB. We tested the tracking performance of the QR method by using a time-varying multipath channel. The simulations show performance of the algorithm for mobiles transmitting so that the maximum Doppler frequency (defined as $f_D = V/\tilde{\lambda}$, where V is vehicle speed and $\tilde{\lambda}$ is carrier wavelength) multiplied by the symbol period T is $f_D T = 0.0006$. Delay intervals are $D_1/T = 0.15,\ D_2/T = 0.35$, and $D_3/T = 0.75$, respectively. The selected parameters are representative of mobiles moving at 27–60 mi/h (depending on the bit rate). In Fig. 8, channel tracking performance of the QR approach on a typical fade for $\mathrm{Re}\{w_2(0)\}$ and SNR $= 30$ dB is shown. The forgetting factor is $\lambda = 0.97$.

The bit-error rate analysis results are shown in Fig. 9 for delay intervals $D_1/T = 0.15,\ D_2/T = 0.2$, and $D_3/T = 0.25$. The channel is time varying with $f_D T = 0.0006$, but during the first 500 symbols, bit errors are not measured to allow convergence of the weights. Channel variations are then tracked by the adaptive filter with a value of $\lambda = 0.98$. The SNR is the same on each discrete-time channel. A sample size of 10^{k+2} was used to estimate an error probability

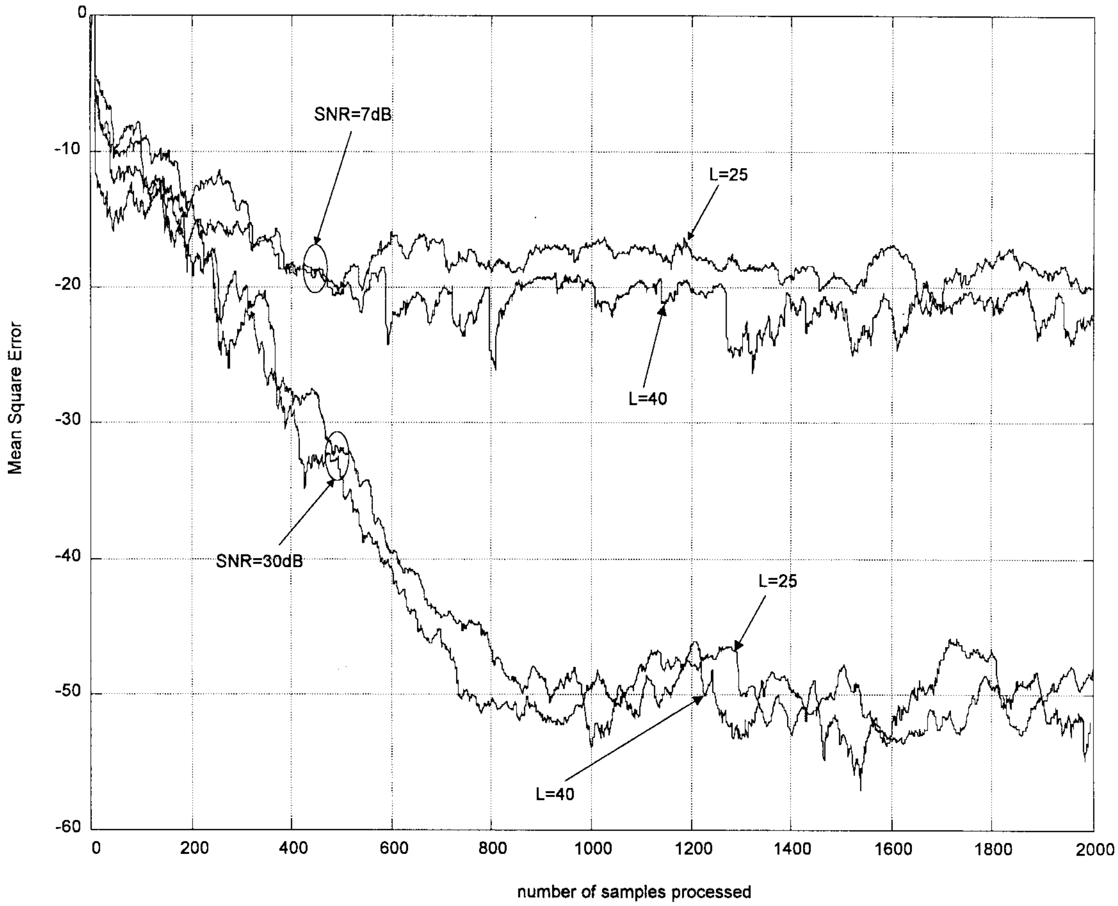


Fig. 5. Convergence process (same conditions as Fig. 4) for different SNR's on the first path—changing the length of the deconvolution baseband filters.

of 10^{-k}. The results of the QR-RLS filter described in Appendix D are also shown as dashed curves. Considerable improvement is achieved as a consequence of the use of HOS. As already pointed out, the HOS method is more efficient in the identification of NMP models, but degradations are expected in the presence of fast fading. There is an evident tradeoff between performance and complexity. We observe that increasing the order of the space-processing architecture (that is, to increase the number of elements in the array) is considerably more expensive than increasing the order of the time-processing architecture (that is, to increase the length of the branch filters). Time processing, in fact, involves only the DSP back end.

VI. Conclusions

A new method has been presented to process digitally modulated signals with an antenna array receiver in a mobile radio environment using high-order statistics. It is based on some of the ideas presented in [3] for the scalar real deconvolution problem. Among the advantages with respect to the minimum mean-square estimator, we can cite the enhanced identification capability (no particular assumptions are required for the channels, particularly the MP property). Certainly, tracking speed is decreased when using high-order statistics because of the larger sample size required to obtain consistent cumulant estimates.

The price for the performance improvement is the increased computational complexity, which may result very high for channels with a large delay spread and large number of sensors. An alternative solution to this problem would be to reduce the size of the linear system and just discarding part of the available equations. The adaptive algorithm is implemented by means of square-root decomposition of the cumulant matrix and has the ability to track time-varying channels. The results of some simulations for a TDMA multipath channel are shown in a single-transmitter environment.

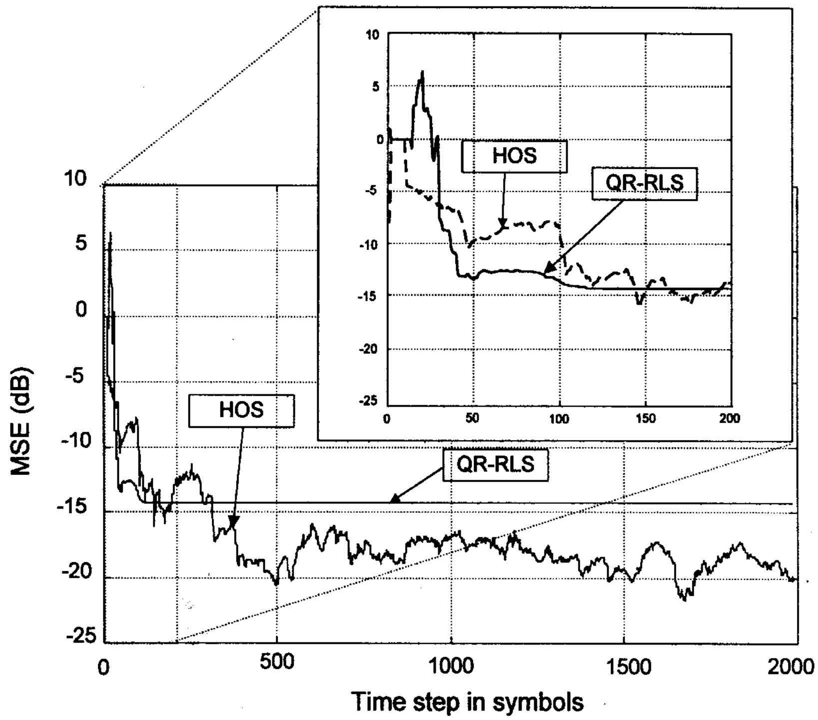
Appendix

A. Cumulants Estimation

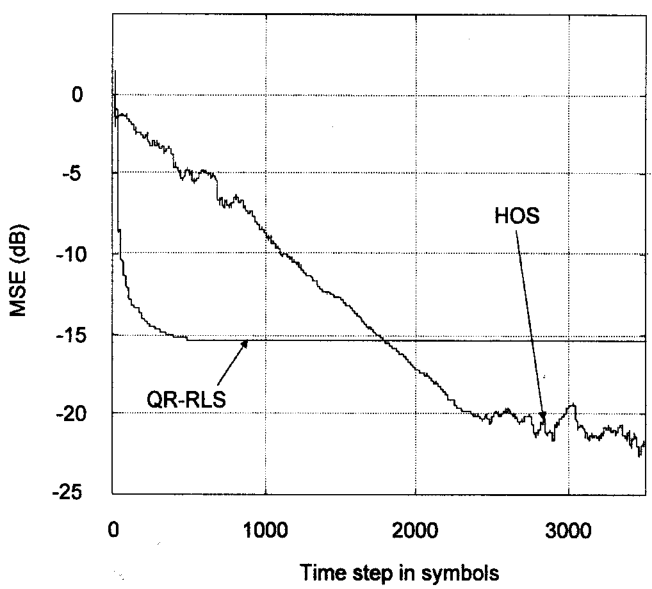
Adaptive estimation of cumulants can be implemented by means of the method also used in [15]. We define

$$\begin{aligned} m^y_{i_1,i_2,i_3,i_4}(k) &= E\{y_{i_1}(n), y^*_{i_2}(n+k), y_{i_3}(n+k) \\ &\quad y^*_{i_4}(n+k)\} \\ m^y_{i,l}(k) &= E\{y_i(n), y^*_l(n+k)\} \end{aligned}$$

and the estimates of the respective moments based on sample statistics as $\hat{m}^y_{i,l_1,l_2,l_3}(k)^{(n)}, \hat{m}^y_{i,l}(k)^{(n)}$ using n samples. Assuming we have available at iteration 0, $y_i(0), \cdots, y_i(I_{lag})$ for



(a)



(b)

Fig. 6. MSE performance for the HOS (blind) algorithm and the QR-RLS (perfectly trained) filter. The delay interval for the three channels are (a) 0.15, 0.25, and 0.5 and (b) 0.15, 0.5, and 0.75.

$i = 1, 2, \cdots, K$, at iteration n we can update $\hat{m}^y_{i,l_1,l_2,l_3,}(k)^{(n)}$ from $\hat{m}^y_{i,l_1,l_2,l_3}(k)^{(n-1)}$ as follows:

$$\begin{aligned}
\hat{m}^y_{i,l_1,l_2,l_3}&(k)^{(n)} \\
&= (1-\alpha(n))\, \hat{m}^y_{i,l_1,l_2,l_3}(k)^{(n-1)} \\
&\quad + \alpha(n) y_i(\Gamma^{(n,k)}) y^*_{l_1}(\Gamma^{(n,k)}+k) y_{l_2}(\Gamma^{(n,k)}+k) \\
&\quad \times y^*_{l_3}(\Gamma^{(n,k)}+k)
\end{aligned}$$

where $\alpha(n) = (1/n) + I_{lag}$ and $\Gamma^{(n,k)} = \min(n + I_{lag}, n + I_{lag} - k)$. Similarly, for the second-order moments, we can write

$$\begin{aligned}
\hat{m}^y_{i,l}(k)^{(n)} &= \hat{c}^y_{i,l}(k)^{(n)} \\
&= (1-\alpha(n))\, \hat{m}^y_{i,l}(k)^{(n-1)} \\
&\quad + \alpha(n) y_i(\Gamma^{(n,k)}) y^*_l(\Gamma^{(n,k)}+k). \qquad (22)
\end{aligned}$$

Evidently, we can obtain the cumulants estimation at point n as

$$\begin{aligned}
\hat{c}^y_{i,l_1,l_2,l_3}&(k)^{(n)} \\
&= \hat{m}^y_{i,l_1,l_2,l_3}(k)^{(n)} \\
&\quad - \hat{m}^y_{i,l_1}(k)^{(n)} \hat{m}^y_{l_2,l_3}(0)^{(n)} - \hat{m}^y_{i,l_2}(k)^{(n)} \hat{m}^y_{l_1,l_3}(0)^{(n)} \\
&\quad - \hat{m}^y_{i,l_3}(k)^{(n)} \hat{m}^y_{l_1,l_2}(0)^{(n)}.
\end{aligned}$$

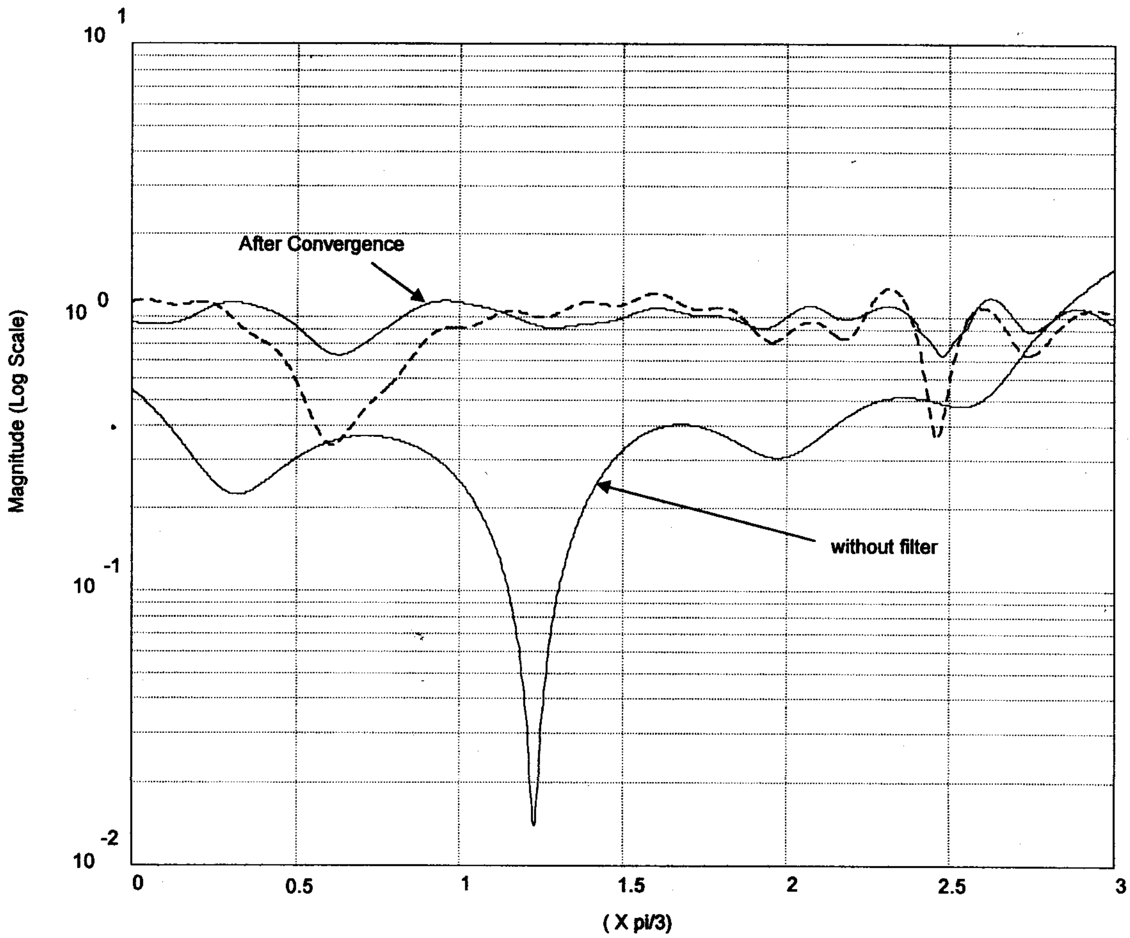


Fig. 7. Frequency response of the discrete-time system (channels-combining scheme) with no filter and after convergence (700 symbols) of the complex weights of the baseband filter (solid). Also, the response after 500 samples (dotted) is provided. The channel is in the same condition as Fig. 4.

Finally, $\hat{s}^{i,l}_{l_1,i_1,k_2}(n)$, the estimate of $s^{i,l}_{l_1,i_1,k_2}$ based on n samples, can be obtained as

$$\begin{aligned}\hat{s}^{i,l}_{l_1,i_1,k_2}(n) = \sum_{m=0}^{q} &[\hat{c}^{y}_{l_1,l}(i_1+m)^{(n)}\hat{c}^{y}_{i,l,l,l}(k_2-m)^{(n)} \\ &- \hat{c}^{y}_{i,l}(k_2-m)^{(n)}\hat{c}^{y}_{l_1,l,l,l}(i_1+m)^{(n)}] \\ &\quad k_2 = -q, -q+1, \cdots, 2q \\ &\quad i, l, l_1 = 1, 2, \cdots, K \\ &\quad i_1 = L_1, \cdots, L_2\end{aligned} \tag{23}$$

and an obvious organization of these quantities into the matrices of the system (19) gives the estimates $\boldsymbol{d}(n)$ and $\boldsymbol{C}(n)$.

B. Proof of (14)

From the cumulant properties, we can write for the following second-order cross moments:

$$\begin{aligned}E\{x(k)y_i(k+n_1)^*\} &= \sum_{m=0}^{q} h_i(m)^* E\{x(k)x(k+n_1-m)^*\} \\ &= \gamma_2 h_i(n_1)^*.\end{aligned} \tag{24}$$

This expression is evidently different from zero only for $0 \le n_1 \le q$. Considering that

$$\begin{aligned}E\{x(k)y_i(k+n_1)^*\} &\simeq E\{z(k)y_i(k+n_1)^*\} \\ &= \sum_{l=1}^{K}\sum_{k_1=L_1}^{L_2} w_l(k_1)E\{y_l(k-k_1)y_i(k+n_1)^*\}\end{aligned} \tag{25}$$

that is,

$$E\{x(k)y_i(k+n_1)^*\} = \sum_{l=1}^{K}\sum_{k_1=L_1}^{L_2} w_l(k_1)c^{y}_{l,i}(k_1+n_1)$$

and equating (25) and (24), we obtain (14).

C. Proof of (18)

Let us consider the z transform of these two sequences

$$\begin{aligned}c^{y}_{i,l}(m) &= \text{cum}\,[y_i(k), y_l^*(k+m)] \\ &= \gamma_2 \sum_{k=0}^{q} h_i(k)h_l(k+m)^*\end{aligned}$$

and

$$c^{y}_{i,l,l,l}(m) = \gamma_4 \sum_{k=0}^{q} h_i(k)h_l(k+m)h_l(k+m)^{*^2}$$

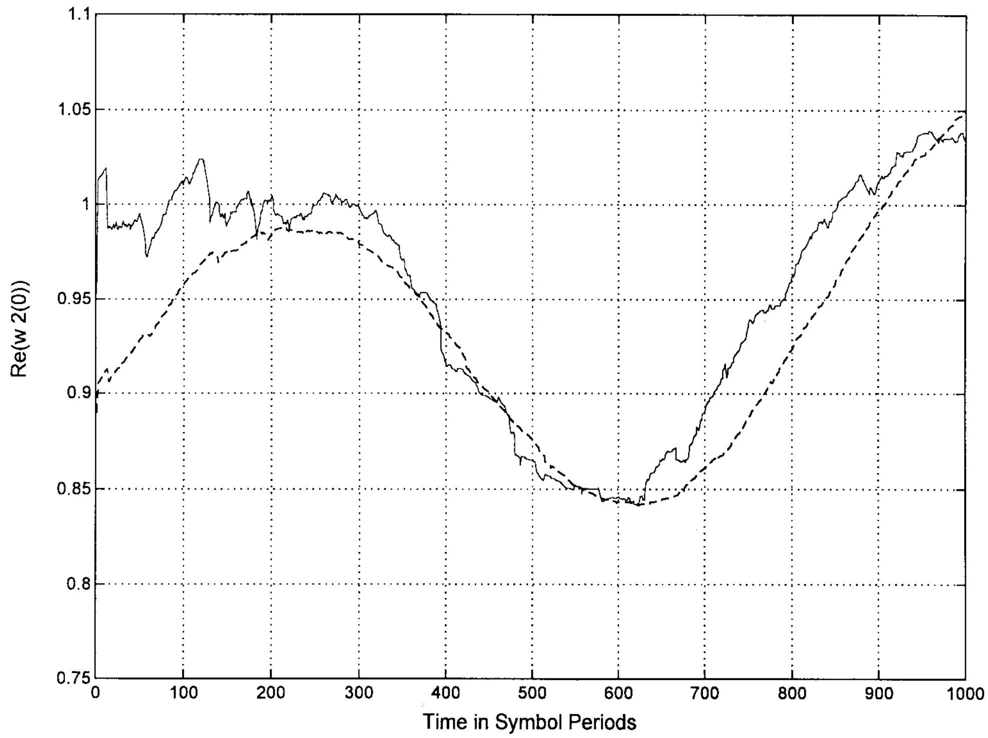


Fig. 8. Tracking performances of the QR approach with equal SNR's on all discrete channels (30 dB). The channel is varying, and the product maximum Doppler frequency-symbol period is equal to 0.0006. Delay intervals are 0.15, 0.35, and 0.75. The forgetting factor is equal to 0.97. The solid curve is the trace of the real part of the weight $w2(0)$ using the adaptive algorithm, and the dashed curve is the optimum performance.

(the z transform of a sequence $b(n)$ is indicated as $\mathcal{Z}[b(n)]$) explicitly

$$\begin{aligned} C_{i,l,l,l}(z) &= \gamma_4 \sum_{m} \sum_{k=0}^{q} h_i(k) h_l(k+m) h_l(k+m)^{*^2} z^{-m} \\ &= \gamma_4 \left(\sum_{\nu_1} h_i(\nu_1) z^{\nu_1} \right) \\ &\quad \cdot \left(\sum_{\nu_2} h_l(\nu_2) h_l(\nu_2)^{*^2} z^{-\nu_2} \right) \\ &= \gamma_4 H_i(z^{-1}) H_l^{(3)}(z) \end{aligned} \tag{26}$$

and

$$C_{i,l}(z) = \gamma_2 H_i(z^{-1}) H_l^*(z)$$

where $H_i(z) = \mathcal{Z}[h_i(n)]$ and $H_l^{(3)}(z) = \mathcal{Z}[h_l(n) h_l(n)^{*^2}]$. Eliminating $H_i(z^{-1})$, we can express the z-domain relationship

$$\gamma_2 C_{i,l,l,l}(z) H_l^*(z) = \gamma_4 C_{i,l}(z) H_l^{(3)}(z)$$

so that in the time domain

$$\begin{aligned} \gamma_2 &\sum_{m=0}^{q} h_l^*(m) c_{i,l,l,l}^y(n-m) \\ &= \gamma_4 \sum_{m=0}^{q} \beta_l(m) c_{i,l}^y(n-m), \qquad i,l = 1,2,\cdots,K \end{aligned} \tag{27}$$

with $\beta_l(m) = h_l(m) h_l(m)^{*^2}$. Substituting (14) for $\gamma_2 h_l(m)^* m = 0,1,\cdots,q$ and (17) for $\gamma_4 \beta_l(m)$, we obtain the set of equations (18).

D. The QR-RLS Adaptive Algorithm

The problem we have to solve at every step is in the case of perfect knowledge of the transmitted sequence $x(n)$

$$\min_{\boldsymbol{w}} \left\| \begin{pmatrix} \lambda \tilde{\boldsymbol{Y}}^{(n)} \\ \tilde{\boldsymbol{y}}(n+1)^T \end{pmatrix} \boldsymbol{w} - \begin{pmatrix} \lambda \boldsymbol{X}^{(n)} \\ x(n+1) \end{pmatrix} \right\|^2 \tag{28}$$

with

$$\begin{aligned} (\tilde{\boldsymbol{Y}}^{(n)})^T &= (\tilde{\boldsymbol{y}}(1), \tilde{\boldsymbol{y}}(2), \cdots, \tilde{\boldsymbol{y}}(n)) \\ \tilde{\boldsymbol{y}}(k) &= (\boldsymbol{y}_1(k), \boldsymbol{y}_2(k), \cdots, \boldsymbol{y}_K(k))^T \\ \boldsymbol{y}_l(k) &= (y_l(k-L_1), y_l(k-L_1+1), \cdots, y_l(L_2))^T \\ \boldsymbol{X}^{(n)} &= (x(1), x(2), \cdots, x(n))^T. \end{aligned}$$

The *normal equations* are now $\tilde{\boldsymbol{Y}}^{(n+1)^{T^*}} \tilde{\boldsymbol{Y}}^{(n+1)} \boldsymbol{w} = \tilde{\boldsymbol{Y}}^{(n+1)^{T^*}} \boldsymbol{X}^{(n+1)}$. Following the same discussion for the HOS case of Section V, with the difference that now the QR decomposition is operated as $\boldsymbol{Q}^H \tilde{\boldsymbol{Y}}^{(n)} = (\binom{\tilde{\boldsymbol{V}}^{(n)}}{0})$ and $\boldsymbol{Q}^H \tilde{\boldsymbol{Y}}^{(n)} = (\binom{\overline{\boldsymbol{X}}^{(n)}}{\overline{x}})$, we can express the algorithm as follows.

1) Compute the prediction error $\tilde{u}(n+1) = x(n+1) - \tilde{\boldsymbol{y}}(n+1)^T \boldsymbol{w}(n)$.
2) Form the matrix

$$\begin{pmatrix} \lambda \tilde{\boldsymbol{V}}^{(n)} & \mathbf{o} \\ \tilde{\boldsymbol{y}}(n+1)^T & \tilde{u}(n+1) \end{pmatrix}.$$

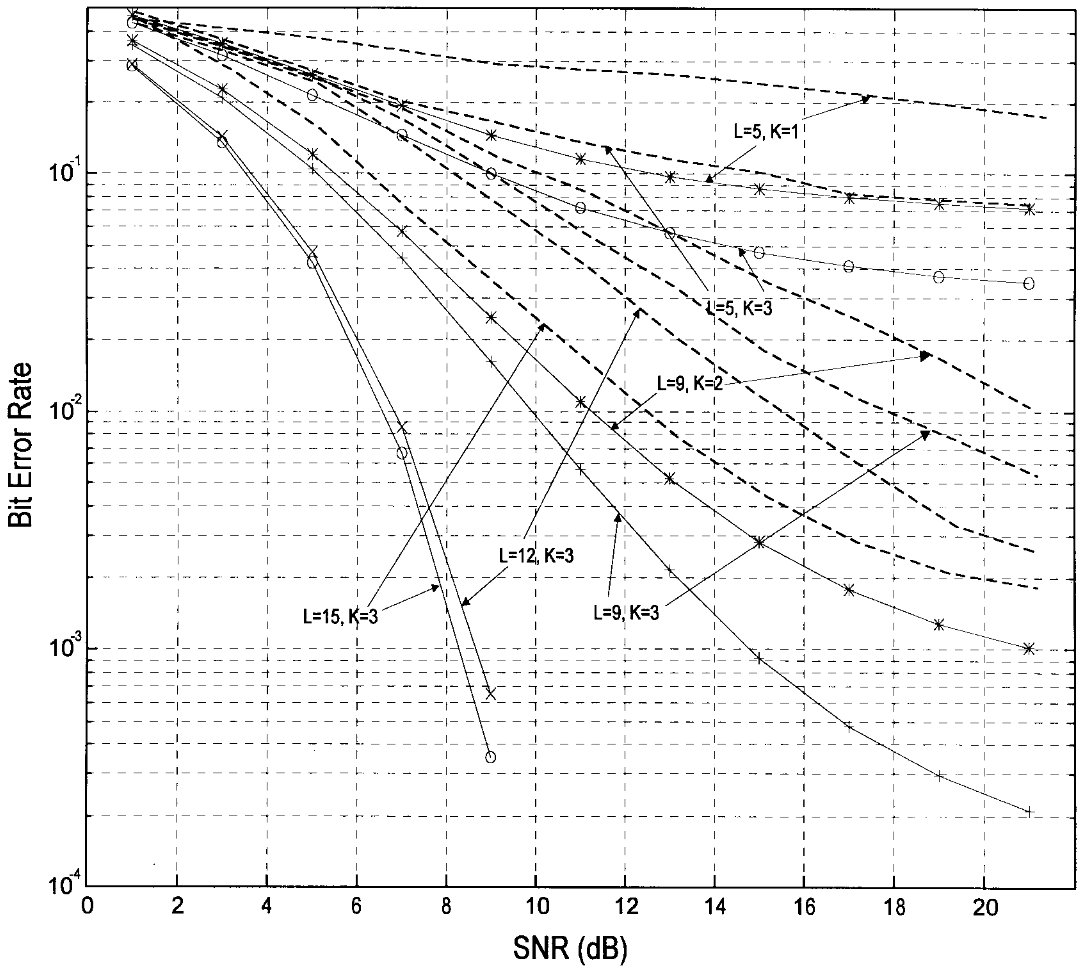


Fig. 9. Bit-error rate performance on a time-varying channel (the product maximum Doppler frequency-symbol period is equal to 0.0006) for different numbers of elements and weights. Delay intervals are 0.15, 0.25, and 0.35. The forgetting factor is 0.98. Solid curves are the performance of the HOS-based method, and dashed curves are the performance of the perfectly trained QR-RLS method.

3) Sweep the bottom part of this matrix using Givens rotations.
4) Solve the triangular system $\tilde{\boldsymbol{V}}^{(n+1)} d\boldsymbol{w}(n+1) = \overline{\boldsymbol{X}}^{(n+1)}$.
5) Obtain $\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + d\boldsymbol{w}(n+1)$.

ACKNOWLEDGMENT

The author would like to acknowledge the positive criticism of the reviewers and many stimulating discussions with A. Faheem who contributed to part of the simulations.

REFERENCES

[1] P. Balaban and J. Saltz, “Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio: Part I, Part II,” *IEEE Trans. Commun.*, vol. 40, pp. 805–907, May 1992.
[2] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.
[3] F. Zheng, S. McLaughlin, and B. Mulgrew, “Cumulant-based deconvolution and identification: Several new families of linear equations,” *Signal Process.*, vol. 30, pp. 199–219, 1993.
[4] B. Porat and B. Friedlander, “FIR system identification using fourth-order cumulants with application to channel equalization,” *IEEE Trans. Automat. Contr.*, vol. 38, no. 9, pp. 1394–1398, 1993.
[5] G. B. Giannakis and J. M. Mendel, “Identification of nonminimum phase systems using higher order statistics,” *Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 360–377, 1989.
[6] J. M. Mendel, “Tutorial on higher order statistics (spectra) in signal processing and system theory: Theoretical results and some applications,” *Proc. IEEE*, vol. 79, pp. 278–305, Mar. 1991.
[7] R. T. Compton, Jr., “The bandwidth performance of a two-element adaptive array with tapped delay-line processing,” *IEEE Trans. Antennas Propagat.*, vol. 36, no. 1, pp. 82–87, 1988.
[8] L. Tong, G. Xu, and T. Kailath, “Fast blind equalization via antenna arrays,” in *Proc. ICASSP ‘93*, vol. IV, pp. 272–275.
[9] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
[10] 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, pp. 1043–1047, Dec. 1978.
[11] P. Balaban and J. Saltz, “Dual diversity combining and equalization in digital cellular mobile radio,” *IEEE Trans. Veh. Technol.*, vol. 40, pp. 342–354, May 1991.
[12] J. H. Winters, “Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading,” *IEEE Trans. Veh. Technol.*, vol. 42, pp. 377–384, Nov. 1993.
[13] ——, “On the capacity of radio communication systems with diversity in a Rayleigh fading environment,” *IEEE J. Select. Areas Commun.*, vol. 2, pp. 579–586, July 1987.
[14] T. Ohgane, T. Shimura, N. Matsuzawa, and H. Sasaoka, “An implementation of a CMA adaptive array for high speed GMSK transmission in mobile communications,” *IEEE Trans. Veh. Technol.*, vol. 42, pp. 282–288, Aug. 1993.

[15] D. Hatzinakos and C. L. Nikias, “Blind equalization using a tricepstrum based algorithm,” *IEEE Trans. Commun.*, vol. 39, pp. 669–682, Aug. 1991.
[16] H. H. Chiang and C. L. Nikias, “Adaptive deconvolution and identification of nonminimum phase FIR systems based on cumulant,” *IEEE Trans. Automat. Contr.*, vol. 35, pp. 36–47, Jan. 1990.
[17] J. F. Bobrow and W. Murray, “An algorithm for RLS identification of parameters that vary quickly with time,” *IEEE Trans. Automat. Contr.*, vol. 38, pp. 351–354, Feb. 1993.
[18] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
[19] M. Rosenblatt, *Stationary Sequences and Random Fields*. Boston, MA: Birkhauser, 1985.
[20] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

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

He worked for the Italian Air Force from 1990 to 1991 and consulted in the area of digital signal processing applied to communications for Staer, Inc., S.P.E., Inc., and 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 Researcher at the ECSE Department, Rensselaer Polytechnic Institute Troy, NY. He was a Wireless Communications Consultant for ATS, Inc., Waltham, MA, from 1994 to 1995. In 1995, he joined the Telecommunications Group of the Watkins-Johnson Company, Gaithersburg, MD, where he is involved in the design of cellular equipment for AMPS and TDMA IS-136. His main interests are in advanced signal processing for HF receivers implementation, spread-spectrum multiple-access communications, and mobile radio communications.

Dr. Martone is a Member of the New York Academy of Sciences, the American Association for the Advancement of Science, and the Associazione Elettrotecnica Italiana.