An Adaptive Algorithm for Antenna Array Low-Rank Processing in Cellular TDMA Base Stations

Massimiliano (Max) Martone, *Member, IEEE*

***Abstract*— A new adaptive algorithm for blind interference rejection and multipath mitigation is studied and applied to antenna array processing in the reverse channel of a time-division multiple-access (TDMA) cellular communication system. The method is based on higher order statistics (HOS) processing of the baseband vector samples at the antenna array output. The similarity between the cumulant-based solution and the standard multivariable least-squares solution is exploited to derive an efficient adaptive algorithm based on a low-rank processing architecture. The algorithm exhibits good tracking and enhanced identification capability with respect to traditional least-squares methods.**

***Index Terms*—Array signal processing, higher order statistics, interference suppression, land mobile radio cellular systems, time-division multiaccess.**

I. Introduction

IN TIME-DIVISION multiple-access (TDMA) systems, data dispersion can span several symbols as a consequence of frequency-selective fading caused by radio-frequency (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 different attenuation and angle of arrival. Space-only processing methods [28], [29] are not effective because intersymbol interference (ISI) cannot be compensated for using the traditional combining architecture. A careful combination of space and time filtering may result in an extremely efficient approach to solve ISI caused by multipath fading and interference caused by multiple cochannel transmitters. The space-time filtering approach results in a discrete-time multiple-input multiple-output (MIMO) model that must be deconvolved. Standard techniques for equalization are based on an equivalent minimum phase system modeling approach because they exploit only the second-order statistics (SOS) of the signal (the minimum mean-square error (MMSE) criterion, for example). However, most real-world channels do not present the minimum phase (MP) condition. Motivated by these observations, the equalization of nonmnimum phase (NMP) channels using higher than second-order statistics (HOS) has stimulated significant interest during the last ten years [22], [17], [14], [8], [25], [19]. Most 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 is not clear. The only blind algorithm well studied in its practical implementation [15] is the constant modulus algorithm (CMA) whose slow convergence under particular situations constitutes the most important objection.

The method proposed in this paper is based on the same idea introduced in [13], where the super-exponential algorithm of [21] was generalized to the multichannel case using third-order cumulants. Here we present the application of the method to the reception of cellular signals using fourth-order cumulants and a new adaptive implementation based on a low-rank processing concept. Low-rank/subspace processing is an extremely important branch of signal processing (see [20] for applications and algorithms). The space/time autocorrelation matrix of signals received at different elements of an array can be ill-conditioned in practice. Any adaptive algorithm based on full-rank processing (for example, traditional recursive least-squares (RLS), [7], [5]) is severly affected by the rank degeneracy problem. We propose here an adaptive algorithm based on a low-rank approximation of the covariance matrix exploiting some important ideas presented by Strobach in [24].

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 result of computer simulations are shown.

II. System Model

We assume U mobile transmitters communicating with a base station with a K-element antenna, with $U \leq K$. The structure of the antenna is assumed to be a uniform linear array, d is the distance between adjacent antenna elements, and $\tilde{\lambda}$ is the wavelength of the signal. Multipath propagation can be characterized for the l-transmitter as an $N^{(l)}$-path channel whose κth path $(\kappa = 1,2,\cdots,N^{(l)})$ is represented by $P_\kappa^{(l)}$ received delayed and attenuated replicas of the signal. The impulse response of the κth path relative to the lth transmitter can be expressed as

$$f_\kappa^{(l)}(t) = \sum_{\nu=1}^{P_\kappa^{(l)}} \rho_{\kappa,\nu}^{(l)} e^{j\psi_{\kappa,\nu}^{(l)}} \delta\left(t - \tau_{\kappa,\nu}^{(l)}\right)$$

Paper approved by S. Ariyavisitakul, the Editor for Wireless Techniques and Fading of the IEEE Communications Society. Manuscript received March 3, 1997; revised September 15, 1997 and December 1, 1997. This paper was presented in part at the IEEE International Conference on Communications, Montreal, P.Q., Canada, June 1997.

The author is with the Telecommunications Group, Watkins-Johnson Company, Gaithersburg, MD 20878-1794 USA (e-mail: max.martone@wj. com).

Publisher Item Identifier S 0090-6778(98)03868-9.

where $\tau_{\kappa,\nu}^{(l)}$, $\rho_{\kappa,\nu}^{(l)}$, and $\psi_{\kappa,\nu}^{(l)}$ are delay, amplitude, and phase of the νth delayed signal in the κth path relative to the lth transmitter, while $\delta(t)$ is the delta function.[1] Here we are assuming a time-invariant channel, whereas $\tau_{\kappa,\nu}^{(l)}$, $\rho_{\kappa,\nu}^{(l)}$, and $\psi_{\kappa,\nu}^{(l)}$ are time-varying parameters. The assumption is justified 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 of the lth transmitter is $m_l(t) = \sum_n x_l(n) p_{tx}(t - nT)$, where $x_l(n) = a_n^{(l)} + j b_n^{(l)}$ are the complex symbols defining the signal constellation used for the particular digital modulation scheme,[2] $p_{tx}(t)$ is a square-root raised-cosine shaping filter with rolloff factor 0.35, and T is the signaling interval. The lth transmitted signal propagated through the κth path can be represented as

$$S_\kappa^{(l)}(t) = \int_{-\infty}^{+\infty} f_\kappa^{(l)}(t-\tau) m_l(\tau)\, d\tau = \sum_{\nu=1}^{P_\kappa^{(l)}} \rho_{\kappa,\nu}^{(l)} e^{j\psi_{\kappa,\nu}^{(l)}} m\big(t - \tau_{\kappa,\nu}^{(l)}\big) e^{j2\pi f_0 (t - \tau_{\kappa,\nu}^{(l)})} \quad (1)$$

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

$$r_k\big(t, \theta_\kappa^{(l)}\big) = e^{j\omega_0 t} \sum_{\nu=1}^{P_\kappa^{(l)}} \rho_{\kappa,\nu}^{(l)} m_l(t - \tau_{\kappa,\nu}) e^{-j2\pi k d \frac{\sin\theta_\kappa^{(l)}}{\lambda}} e^{j\phi_{\kappa,\nu}^{(l)}}$$

where $\phi_{\kappa,\nu}^{(l)} = -2\pi f_0 \tau_{\kappa,\nu}^{(l)} + \psi_{\kappa,\nu}^{(l)}$. 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

$$y_k(n) = \sum_{l=1}^{U} \sum_m h_{k,l}(m) x_l(n-m) + \eta_k(n), \qquad k = 1, 2, \cdots, K \quad (2)$$

where $\eta_k(n)$ is Gaussian white noise and $h_{k,l}(m)$ is the T-sampled[3] impulse response

$$h_{k,l}(m) = \sum_{\kappa=1}^{N_l} \sum_{\nu=1}^{P_\kappa^{(l)}} \rho_{\kappa,\nu}^{(l)} r_p\big(t - \tau_{\kappa,\nu}^{(l)}\big) e^{j\left[\phi_{\kappa,\nu}^{(l)} - 2\pi k d \frac{\sin\theta_\kappa^{(l)}}{\lambda}\right]} \Bigg|_{t=mT} \quad (3)$$

In this expression $r_p(t)$ is the raised-cosine function with excess bandwidth 0.35 [18] obtained because we assume that the receiver filters $p_{rx}(t)$ at each antenna element are square-root raised-cosine filters perfectly matched to the transmitter filters $p_{tx}(t)$. In the following derivation vectors and matrices are bold. $\mathbf{M}^T$, $\mathbf{v}^T$, $\mathbf{M}^H$, and $\mathbf{v}^H$ designate transposition and Hermitian for matrix $\mathbf{M}$ and vector $\mathbf{v}$, respectively. Complex conjugation for scalars, matrices, and vectors is indicated as u^*, $\mathbf{M}^*$, and $\mathbf{v}^*$, 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. We use the notation $\|\mathbf{v}\| = \sqrt{\sum_{i=1}^{M} |v_i|^2}$ for the two-norm of the complex M-vector $\mathbf{v} = [v_1, \cdots, v_M]^T$, and $\|\mathbf{M}\|_F = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} |m_{i,j}|^2}$ for the Frobenius norm of the $M \times N$ complex matrix $\mathbf{M}$ whose generic i,j element is $m_{i,j}$.

We will not consider the contribution of the additive noise in the derivation of the deconvolution algorithm. In the z-domain the transfer function (2) can be expressed as

$$\tilde{\mathcal{H}}(z) = \sum_{m=-\infty}^{\infty} \mathbf{H}(m) z^{-m} \quad (4)$$

where the organization of the $\mathcal{H}_{i,l}(z)$ polynomials [z-transforms of $h_{i,l}(m)$] in $\tilde{\mathcal{H}}(z)$ is given by

$$\tilde{\mathcal{H}}(z) = \begin{bmatrix} \mathcal{H}_{1,1}(z) & \cdots & \mathcal{H}_{1,U}(z) \\ \cdots & \cdots & \cdots \\ \mathcal{H}_{K,1}(z) & \cdots & \mathcal{H}_{K,U}(z) \end{bmatrix}.$$

A. Distortionless Reception

To recover the input signals, a linear K-input U-output filter $\tilde{\mathcal{W}}(z) = \sum_{m=L_1}^{L_2} \mathbf{W}(m) z^{-m}$ with length $L = L_2 - L_1 + 1$ is applied to the outputs of the sensors. The objective for $\tilde{\mathcal{W}}(z)$ is to achieve *distortionless reception.*
If we define

$$\tilde{\mathcal{W}}(z) = \begin{bmatrix} \mathcal{W}_{1,1}(z) & \cdots & \mathcal{W}_{1,K}(z) \\ \cdots & \cdots & \cdots \\ \mathcal{W}_{U,1}(z) & \cdots & \mathcal{W}_{U,K}(z) \end{bmatrix}$$

distortionless reception means that

$$\tilde{\mathcal{W}}(z)\tilde{\mathcal{H}}(z) = \mathbf{I}_U \quad (5)$$

where $\mathbf{I}_U$ is a $U \times U$ identity matrix [9], [4]. 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 signals $x_i(n)$ reflect the same symmetry. Moreover, *signal*

[1] In this model the κth path for the lth transmitter consists of $P_\kappa^{(l)}$ delayed replicas of the signal with the same angle of arrival due to the scatterers nearby the mobile. In fact, assuming the scatterers evenly spread out on a circle surrounding each mobile, and assuming large distance between the mobile and the base station, simple geometric considerations [1] can lead to the simplification of a *point-source approximation* for the scattering mechanism local to the the mobile; that is, we can assume that $P_\kappa^{(l)}$ delayed replicas of the signal are received with approximately the same angle of arrival. For a certain number $N^{(l)}$ of reflections of the lth transmitted signal, particularly reflections in the vicinity of the base station, these assumptions are not reasonable and different angles of arrival have to be considered.

[2] In $\pi/4$ DQPSK [26], [27] we have $\bar{a}_m^{(l)} = \bar{a}_{m-1}^{(l)} \cos[\Delta\phi_m^{(l)}] - \bar{b}_{m-1}^{(l)} \sin[\Delta\phi_m^{(l)}]$, $\bar{b}_m^{(l)} = \bar{a}_{m-1}^{(l)} \sin[\Delta\phi_m^{(l)}] + \bar{b}_{m-1}^{(l)} \cos[\Delta\phi_m^{(l)}]$, where $\Delta\phi_m^{(l)} = \pi/4$ if $\mathrm{bit}_{1,\mathrm{m}}^{(\mathrm{l})} = 0$ and $\mathrm{bit}_{2,\mathrm{m}}^{(\mathrm{l})} = 0$; $\Delta\phi_m^{(l)} = 3\pi/4$ if $\mathrm{bit}_{1,\mathrm{m}}^{(\mathrm{l})} = 1$ and $\mathrm{bit}_{2,\mathrm{m}}^{(\mathrm{l})} = 0$; $\Delta\phi_m = -3\pi/4$ if $\mathrm{bit}_{1,\mathrm{m}}^{(\mathrm{l})} = 1$ and $\mathrm{bit}_{2,\mathrm{m}}^{(\mathrm{l})} = 1$; and $\Delta\phi_m = -\pi/4$ if $\mathrm{bit}_{1,\mathrm{m}}^{(\mathrm{l})} = 0$ and $\mathrm{bit}_{2,\mathrm{m}}^{(\mathrm{l})} = 1$.

[3] The same model with some marginal changes applies to the fractional sampling case as well. We restrict, however, the description of the algorithm to the symbol-spaced case for the sake of clarity.

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. The best possible result for practical *distortionless reception* by means of a linear filter is

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

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

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

where $\phi_i \in [-\pi, \pi]$, n_i is an integer for $i = 1, 2, \cdots, U$ [25], [10], [11]. We say that $\tilde{\mathcal{H}}(z)$ satisfies the *distortionless reception* condition if there exists a BIBO stable *distortionless reception* filter $\tilde{\mathcal{W}}(z)$. A system $\tilde{\mathcal{H}}(z)$ satisfies the distortionless reception condition if and only if [10], [11]

$$\det(\tilde{\mathcal{H}}^H(e^{j\omega})\tilde{\mathcal{H}}(e^{j\omega})) \neq 0, \qquad \text{for all } \omega \in [-\pi, \pi]. \tag{7}$$

B. Vector Organization of Impulse Responses

In the time domain the linear filter can be written

$$z_i(n) = \sum_{j=1}^{K}\sum_{m} w_{i,j}(m)y_j(n-m), \qquad i = 1, 2, \cdots, U \tag{8}$$

where $w_{i,j}(m)$ is the filter corresponding to the polynomial $\mathcal{W}_{i,j}(z)$ and $z_i(n)$ is the ith output of the deconvolution filter $\tilde{\mathcal{W}}(z)$. The overall impulse response is characterized by the input/output relation

$$z_i(n) = \sum_{j=1}^{U}\sum_{m} s_{i,j}(m)x_j(n-m), \qquad i = 1, 2, \cdots, U. \tag{9}$$

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

$$\begin{aligned} s_{i,j_1}(m_2) &= \sum_{j=1}^{K}\sum_{m} w_{i,j}(m)h_{j,j_1}(m_2 - m) \\ &= \left[\sum_{j=1}^{K} \mathbf{H}_{j,j_1}\mathbf{w}_{i,j}\right]_{m_2}, \qquad i, j_1 = 1, 2, \cdots, U \end{aligned}$$

where

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

and $\mathbf{w}_{i,j} = [w_{i,j}(L_1), w_{i,j}(L_1+1), \cdots, w_{i,j}(L_2)]^T$. In a vector form we can write

$$\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i, \qquad i = 1, 2, \cdots, U \tag{10}$$

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

$$\begin{aligned} \tilde{\mathbf{H}} &= \begin{bmatrix} \mathbf{H}_{1,1} & \mathbf{H}_{2,1} & \cdots & \mathbf{H}_{K,1} \\ \mathbf{H}_{1,2} & \mathbf{H}_{2,2} & \cdots & \mathbf{H}_{K,2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{H}_{1,U} & \mathbf{H}_{2,U} & \cdots & \mathbf{H}_{K,U} \end{bmatrix} \\ \tilde{\mathbf{w}}_i &= \left[\mathbf{w}_{i,1}^T, \mathbf{w}_{i,2}^T, \cdots, \mathbf{w}_{i,K}^T\right]^T \\ \tilde{\mathbf{s}}_i &= \left[\mathbf{s}_{i,1}^T, \mathbf{s}_{i,2}^T, \mathbf{s}_{i,3}^T, \cdots, \mathbf{s}_{i,U}^T\right]^T \\ \mathbf{s}_{i,j} &= \left[\cdots, s_{i,j}(-1), s_{i,j}(0), s_{i,j}(1), \cdots\right]^T. \end{aligned}$$

The desired response $\tilde{\mathbf{s}}_i$ that completely restores the information signal of the ith transmitter up to the delay n_i can be expressed as

$$\tilde{\boldsymbol{\delta}}_i = \left[\boldsymbol{\delta}_{i,1}^T, \boldsymbol{\delta}_{i,2}^T, \boldsymbol{\delta}_{i,3}^T, \cdots, \boldsymbol{\delta}_{i,U}^T\right]^T \tag{11}$$

where

$$\boldsymbol{\delta}_{i,j} = \begin{cases} [\cdots, 0, 0, 0, \cdots], & i \neq j \\ \hat{\boldsymbol{\delta}}_i, & i = j. \end{cases} \tag{12}$$

The generic mth element of the vector $\hat{\boldsymbol{\delta}}_i$ is $\delta(m - n_i)$, if we neglect the phase shift and we force the solution not to permute the inputs ($\mathbf{P} = \mathbf{I}_U$).[4] This generalizes the one-dimensional case [21]. It is possible to solve this deconvolution problem by solving the minimization problem

$$\min_{\tilde{\mathbf{w}}_i} ||\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i - \tilde{\boldsymbol{\delta}}_i||^2, \qquad i = 1, 2, \cdots, U. \tag{13}$$

In the following sections we will solve the problem of finding the filter $\tilde{\mathbf{w}}_i$ that restores only the ith transmitted digital stream, which is the signal of interest. That is why we omit the notation $i = 1, 2, \cdots, U$ in the following expressions.

C. Key Assumptions and Their Justification

We make now the important assumptions on the discrete-time model described and we detail some requirements on the statistical properties of the input symbols distribution. The HOS 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 [14]. The properties of cumulants that we exploit are:

- LIN: $\text{cum}[\sum_i \text{f(i)x(i)}, \cdots] = \sum_i \text{f(i)}\,\text{cum}[\text{x(i)}, \cdots]$;
- 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:

- **AS1**: $\tilde{\mathcal{H}}(z)$ is *irreducible* and $\mathbf{H}(m) \neq \mathbf{0}$ only for $m \in [J_1, J_2]$ ($J = J_2 - J_1 + 1$) with $\mathbf{H}(J_2)$ full rank;
- **AS2**: 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:
  - $E\{x_i(n)\} = E\{x_i^3(n)\} = 0$
  - $\text{cum}[\text{x}_\text{i}\text{(n)}, \text{x}_\text{j}^*\text{(n)}] = \sigma_\text{x}^2 > 0$, only for $i = j$
  - $E\{x_i(n)x_j(n)\} = E\{x_i^*(n)x_j^*(n)\} = 0$, for any i, j
  - $\text{cum}[\text{x}_{\text{j}_1}\text{(n)}, \text{x}_{\text{j}_2}\text{(n)}, \text{x}_{\text{j}_3}^*\text{(n)}, \text{x}_{\text{j}_4}^*\text{(n)}] = \gamma_{4\text{x}} \neq 0$, only for $j_1 = j_2 = j_3 = j_4$.

[4] In practice the permutation ambiguity is not a problem because it can be solved by a user identification matching procedure. Specifically, in [26] and [27] the coded digital verification color code (CDVCC) section of the frame is intended to be used for this purpose.

Assumption **AS1** is required to ensure the *distortionless reception* condition for $\tilde{\mathcal{H}}(z)$. In fact, under **AS1** it is possible to design a deconvolution filter with length $L = L_2 - L_1 + 1 \geq \lceil ((J-1)U)/(K-U)) \rceil$ such that the *generalized Sylvester matrix* $H_{L,J}$ of the MIMO system is full column rank [9], [4]. This implies that: 1) a *distortionless reception* linear filter exists (i.e., $\tilde{\mathcal{H}}(z)$ satisfies the *distortionless reception* condition) and 2) (7) is verified. Assumption **AS2** requires the same zero-lag fourth-order cumulant and variance for the U modulated signal sources $x_i(n)$. Since the U signals are assumed to use the same modulation scheme, this appears reasonable. However, a generalization of the algorithm to the case of different zero-lag fourth-order cumulants and variances is possible. In practical applications when $\tilde{\mathcal{H}}(z)$ is a time-varying response, it is almost impossible to guarantee at any time instant the existence of the full-rank least-squares solution of (13).[5] In fact, even if some propagation parameters may be considered almost time-invariant (for example, delay spread or maximum Doppler shift), the characteristics of the different multipath channels are independently changing in time and it is, in general, quite unrealistic to make such a strong assumption. The well-known remedy to rank-deficient (or ill-conditioned) least-squares problems is the use of a minimum-norm solution of (13) based on a low-rank approximation of the Hermitian matrix $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$, which is obtained using a dominant eigendecomposition.[6] This idea will be later developed in an adaptive algorithm based on a subspace tracking concept. *The low-rank approximation can also be motivated by computational efficiency as it will be clear in Section IV.* The eigenvalue decomposition of $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$ is $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}} = \tilde{\mathbf{V}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{V}}^H$, where $\tilde{\mathbf{V}} = [\tilde{\mathbf{v}}_1 \ \tilde{\mathbf{v}}_2 \ \cdots \ \tilde{\mathbf{v}}_{KL}]$ is the $KL \times KL$ orthonormal matrix of eigenvectors, and $\tilde{\mathbf{\Lambda}} = \mathrm{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, \cdots, \tilde{\lambda}_{KL})$ is the diagonal matrix of eigenvalues. The r dominant eigenvalues are those eigenvalues that exceed a certain predetermined tolerance level tol. Hence, using the notation $\mathbf{V}_n = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n]$, $\mathbf{\Lambda}_n = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, we can write $\tilde{\mathbf{V}} = [\tilde{\mathbf{V}}_r \ \tilde{\mathbf{V}}_{KL-r}]$ and $\tilde{\mathbf{\Lambda}} = \begin{bmatrix} \tilde{\mathbf{\Lambda}}_r & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{\Lambda}}_{KL-r} \end{bmatrix}$. In practice one has to find a sufficiently large integer r such that $\tilde{\mathbf{\Lambda}}_r^{-1}$ exists. Next we define the Moore–Penrose inverse $(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^\dagger = \tilde{\mathbf{V}}_r\tilde{\mathbf{\Lambda}}_r^{-1}\tilde{\mathbf{V}}_r^H$ so that we can solve the least-squares problem (13) as

$$\tilde{\mathbf{w}}_i = (\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^\dagger\tilde{\mathbf{H}}^H\tilde{\boldsymbol{\delta}}_i = \tilde{\mathbf{V}}_r\tilde{\mathbf{\Lambda}}_r^{-1}\tilde{\mathbf{V}}_r^H\tilde{\mathbf{H}}^H\tilde{\boldsymbol{\delta}}_i. \tag{14}$$

[5] We are trying to point out that the Hermitian matrix $\tilde{\mathbf{H}}^H\tilde{\mathbf{H}}$ may be not only singular but also *close* to singular—that happens when $\tilde{\mathbf{H}}$ is nearly rank-deficient; in other words, the least-squares problem (13) is *ill conditioned*.

[6] It is interesting to observe that this modification to the super-exponential algorithm was proposed independently by Ding in [6], where it was shown that the *length and zero condition* in addition to the low-rank modification results in a globally convergent algorithm. Observe that the case treated in [6] is the single-input multiple-output (SIMO) case, and that we are instead treating the more general MIMO case—in fact, the *length and zero condition* translates into **AS1**. It is also important to note that [6] does not address the problem of extreme practical interest of designing an adaptive algorithm to implement the batch low-rank super-exponential procedure. This is exactly the problem we are going to solve in the following sections.

III. Derivation of the Algorithm

The following two-step iterative procedure defines a class of algorithms for different values of p and q [21], [13]:

$$s_{i,j}(k)^{[1]} = (s_{i,j}(k))^p(s_{i,j}(k)^*)^q \tag{15}$$

$$s_{i,j}(k)^{[2]} = s_{i,j}(k)^{[1]}\frac{1}{\sqrt{\sum_{j=1}^{K}\sum_l |s_{i,j}(l)^{[1]}|^2}} \tag{16}$$

where $(\cdot)^{[1]}$ and $(\cdot)^{[2]}$ stand for the result of the first and second step, respectively. This iterative algorithm converges at a "*super-exponential*" rate to the desired solution [21] for $p + q \geq 2$. In this work 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 [13] in terms of $\tilde{\mathbf{w}}_i$. If we define $\tilde{\mathbf{g}}_i = [\mathbf{g}_{i,1}^T, \mathbf{g}_{i,2}^T, \mathbf{g}_{i,3}^T, \cdots, \mathbf{g}_{i,U}^T]^T$, with $\mathbf{g}_{i,j} = [\cdots, g_{i,j}(-1), g_{i,j}(0), g_{i,j}(1), \cdots]^T$ as the impulse response vector obtained by $g_{i,j}(k) = s_{i,j}^2(k)s_{i,j}(k)^*$, we can state the least-squares minimization problem

$$\min_{\tilde{\mathbf{w}}_i} \left\| \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i^{[1]} - \tilde{\mathbf{g}}_i \right\|^2 \tag{17}$$

whose solution is

$$\tilde{\mathbf{w}}_i^{[1]} = (\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^\dagger\tilde{\mathbf{H}}^H\tilde{\mathbf{g}}_i. \tag{18}$$

To obtain normalization (16), the second step is

$$\tilde{\mathbf{w}}_i^{[2]} = \frac{\tilde{\mathbf{w}}_i^{[1]}}{\sqrt{\tilde{\mathbf{w}}_i^{[1]^H}(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})\tilde{\mathbf{w}}_i^{[1]}}}. \tag{19}$$

It should be evident that the solution of the least-squares problem (17) [given by (18)] and subsequent normalization as in (19) is equivalent to one step of the iterative procedure (15), (16). Since (15) and (16) converge to the desired solution (11), (12) expressed in terms of $\tilde{\mathbf{w}}_i$ in (13), it is also true that at the point of convergence the solution of (17) is equivalent to the solution of (13). A discussion on the convergence of (18) and (19), related to the convergence of (15) and (16), is given in the Appendix. The procedure (18), (19) can be expressed in terms of the cumulants of the outputs of the sensors. From $y_i(k-n) = \sum_{j=1}^{K}\sum_m h_{i,j}(m-n)x_j(k-m)$, we exploit **AS2** and the properties of the cumulants of linear stationary processes (see property LIN) so that we can write

$$\begin{aligned}
&\mathrm{cum}\left[\mathrm{y_{i_1}(k-n)}, \mathrm{y_{i_2}^*(k-m)}\right] \\
&\quad = \mathrm{cum}\left[\sum_{\mathrm{j_1}=1}^{\mathrm{U}}\sum_{\mathrm{m_1}} \mathrm{h_{i_1,j_1}(m_1-n)x_{j_1}(k-m_1)}, \right. \\
&\qquad \left. \sum_{j_2=1}^{U}\sum_{m_2} h_{i_2,j_2}^*(m_2-m)x_{j_2}^*(k-m_2)\right] \\
&\quad = \sum_{j=1}^{U}\sum_k h_{i_1,j}(k-n)h_{i_2,j}^*(k-m)\sigma_x^2 \\
&\quad = \sigma_x^2\sum_{j=1}^{U}\left[\mathbf{H}_{i_2,j}^H\mathbf{H}_{i_1,j}\right]_{m,n}
\end{aligned} \tag{20}$$

due to

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

To derive the second key expression related to the solution (18), let us consider

$$\mathrm{cum}[\mathrm{z_i(k)}, \mathrm{z_i(k)}, \mathrm{z^*_i(k)}, \mathrm{y^*_{i_3}(k-n)}] = \sum_{j=1}^{U} \sum_{m} h^*_{i_3,j}(m-n)\, \mathrm{cum}[\mathrm{z_i(k)}, \mathrm{z_i(k)}, \mathrm{z^*_i(k)}, \mathrm{x^*_j(k-m)}]$$

and[7]

$$\begin{aligned} &\mathrm{cum}[z_i(k), z_i(k), z_i^*(k), x_j^*(k-m)] \\ &\quad = \sum_{j_1=1}^{U} \sum_{j_2=1}^{U} \sum_{j_3=1}^{U} \sum_{l_1} \sum_{m_1} \sum_{n_1} s_{i,j_1}(l_1) s_{i,j_2}(m_1) s^*_{i,j_3}(n_1) \\ &\qquad \cdot \mathrm{cum}[x_{j_1}(k-l_1), x_{j_2}(k-m_1), \\ &\qquad\qquad x^*_{j_3}(k-n_1), x_j^*(k-m)] \\ &\quad = \gamma_{4x} s^2_{i,j}(m) s^*_{i,j}(m), \end{aligned}$$

So we can write

$$\begin{aligned} \mathrm{cum}\big[\mathrm{z_i(k)}, \mathrm{z_i(k)}, \mathrm{z^*_i(k)}, \mathrm{y^*_{i_3}(k-n)}\big] &= \sum_{j=1}^{U} \sum_{m} h^*_{i_3,j}(m-n) g_{i,j}(m) \gamma_{4x} \\ &= \gamma_{4x} \sum_{j=1}^{U} \big[\mathbf{H}^H_{i_3,j} \mathbf{g}_{i,j}\big]_n. \end{aligned} \tag{21}$$

Expressions (20) and (21) can be substituted in the least-squares solution (18) and the following iterative algorithm is obtained:

$$\begin{aligned} \tilde{\mathbf{w}}_i^{[1]} &= \tilde{\mathbf{\Phi}}^{\dagger} \mathbf{D}_i \\ \tilde{\mathbf{w}}_i^{[2]} &= \frac{\tilde{\mathbf{w}}_i^{[1]}}{\sqrt{\tilde{\mathbf{w}}_i^{[1]^H} \tilde{\mathbf{\Phi}} \tilde{\mathbf{w}}_i^{[1]}}} \end{aligned} \tag{22}$$

where the matrix $\tilde{\mathbf{\Phi}}$ with dimensions $KL \times KL$ is given by

$$\tilde{\mathbf{\Phi}} = \begin{bmatrix} \mathbf{\Phi}_{1,1} & \mathbf{\Phi}_{1,2} & \cdots & \mathbf{\Phi}_{1,K} \\ \mathbf{\Phi}_{2,1} & \mathbf{\Phi}_{2,2} & \cdots & \mathbf{\Phi}_{2,K} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{\Phi}_{K,1} & \mathbf{\Phi}_{K,2} & \cdots & \mathbf{\Phi}_{K,K} \end{bmatrix} \tag{23}$$

$$[\mathbf{\Phi}_{i,j}]_{n,m} = \frac{\mathrm{cum}[\mathrm{y_j(k-m)}, \mathrm{y^*_i(k-n)}]}{\sigma^2_{\mathrm{x}}} \tag{24}$$

[7] Due to **AS2**

$$\mathrm{cum}\big[x_{j_1}(n-m_1), x_{j_2}(n-m_2), x^*_{j_3}(n-m_3), x^*_j(n-k)\big] = \begin{cases} \gamma_{4x}, & j_1 = j_2 = j_3 = j;\ m_1 = m_2 = m_3 = k \\ 0, & \text{otherwise.} \end{cases}$$

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

$$\mathbf{D}_i = \big[\mathbf{d}^T_{i,1}, \mathbf{d}^T_{i,2}, \cdots, \mathbf{d}^T_{i,K}\big]^T \tag{25}$$

$$[\mathbf{d}_{i,m}]_n = \frac{\mathrm{cum}[z_i(k), z_i(k), z_i^*(k), y_m^*(k-n)]}{\gamma_{4x}}. \tag{26}$$

IV. Adaptive Low-Rank Processing

In this section we derive an adaptive algorithm for online computation of the multichannel deconvolution filter parameters. The cumulants of interest in the algorithm can be estimated as

$$\overline{\mathrm{cum}}[y(k-n), y^*(k-m)] = \frac{1}{N} \sum_{k=1}^{N} y(k-n) y^*(k-m) \tag{27}$$

$$\begin{aligned} \overline{\mathrm{cum}}[z(k), z(k), z^*(k), y^*(k-n)] &= \frac{1}{N} \sum_{k=1}^{N} |z(k)|^2 y^*(k-n) z(k) \\ &\quad - 2 \frac{1}{N} \sum_{k=1}^{N} |z(k)|^2 \frac{1}{N} \sum_{k=1}^{N} z(k) y^*(k-n) \\ &\quad - \frac{1}{N} \sum_{k=1}^{N} z(k)^2 \frac{1}{N} \sum_{k=1}^{N} z^*(k) y^*(k-n) \end{aligned} \tag{28}$$

where we have neglected indexes for $y(k)$, $z(k)$ for simplicity, and we have indicated the estimated cumulant as $\overline{\mathrm{cum}}(\cdot\cdot)$. At the end of the convergence process

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

must be satisfied. We assume that the power constraint (19) is always satisfied using an automatic gain control (AGC). Clearly at each $n+1$ stage we wish to solve the problem

$$\min_{\tilde{\mathbf{w}}_i} \left\| \begin{bmatrix} \lambda \tilde{\mathbf{Y}}^{(n)} \\ \tilde{\mathbf{y}}(n+1)^T \end{bmatrix} \tilde{\mathbf{w}}_i - \begin{bmatrix} \lambda \tilde{\mathbf{Z}}_i^{(n)} \\ \tilde{z}_i(n+1) \end{bmatrix} \right\|^2 \tag{29}$$

with $\tilde{\mathbf{Y}}^{(n)} = \left[\begin{smallmatrix} \lambda \tilde{\mathbf{Y}}^{(n-1)} \\ \tilde{\mathbf{y}}(n)^T \end{smallmatrix}\right]$, $\tilde{\mathbf{Z}}_i^{(n)} = \left[\begin{smallmatrix} \lambda \tilde{\mathbf{Z}}_i^{(n-1)} \\ \tilde{z}_i(n) \end{smallmatrix}\right]$, $\tilde{\mathbf{y}}(n) = [\breve{\mathbf{y}}_1(n)^T, \breve{\mathbf{y}}_2(n)^T, \cdots, \breve{\mathbf{y}}_K(n)^T]^T$, $\breve{\mathbf{y}}_i(n) = [y_i(n-L_1), y_i(n-L_1-1), \cdots, y_i(n-L_2)]^T$, and $\tilde{z}_i(n) = (|z_i(n-1)|^2 - 2\sigma_x^2) z_i(n-1)$. In fact, the *normal equations* define the desired minimizer as $\tilde{\mathbf{Y}}^{(n+1)^H} \tilde{\mathbf{Y}}^{(n+1)} \tilde{\mathbf{w}}_i = \tilde{\mathbf{Y}}^{(n+1)^H} \tilde{\mathbf{Z}}_i^{(n+1)}$, which is equivalent to $\tilde{\mathbf{\Phi}} \tilde{\mathbf{w}}_i = \mathbf{D}_i$ for $n \to \infty$ and if we employ sample statistics estimators for the cumulants and the covariance matrix. The expression for $\tilde{z}_i(n)$ can be justified by considering the estimation of fourth-order cumulants based on the sample average given by (28) and **AS2** (the third term on the right-hand side of (28) needs not to be estimated since $E\{z_i^2(n)\} = \sum_j \sum_k s_{i,j}(k)^2 E\{x_i(n) x_i(n)\} = 0$, see **AS2**). Moreover, due to the power normalization (19)

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

Observe that the expression (29) reveals the similarity with the RLS problem where the process $\tilde{z}_i(n)$ is the desired signal. The adaptive solution of the problem could be derived using the RLS algorithm or its variations [7], [5], [18]. One of the important drawbacks of full-rank RLS is that this approach requires at every step, explicitly or implicitly, the existence of the inverse $(\tilde{\mathbf{Y}}^{(n+1)^H}\tilde{\mathbf{Y}}^{(n+1)})^{-1}$. This cannot realistically be guaranteed in any practical situation and, in particular, in space/time applications. In fact, the multipath channels follow independent time-varying characteristics so that the data matrix $\tilde{\mathbf{Y}}^{(n)}$ may become ill conditioned due to statistical fluctuation of the received signals. The most immediate consequence is that rank degeneracy appears in the least-squares problem (29), which makes the results of any traditional numerical algorithm based on full-rank processing meaningless. In this case one should detect the rank degeneracy by looking at the singular values of $\tilde{\mathbf{Y}}^{(n)}$ and decide that certain columns of this matrix can be ignored. In particular, the $KL-r$ eigenvectors of $\tilde{\mathbf{Y}}^{(n)^H}\tilde{\mathbf{Y}}^{(n)}$ (the left singular vectors of $\tilde{\mathbf{Y}}^{(n)}$) corresponding to the "small" singular values can be ignored in a minimum-norm solution of the least-squares problem (29).

The eigenvalue decomposition of $\tilde{\mathbf{Y}}^{(n)^H}\tilde{\mathbf{Y}}^{(n)}$ is

$$\tilde{\mathbf{Y}}^{(n)^H}\tilde{\mathbf{Y}}^{(n)} = \mathbf{V}(n)\mathbf{\Lambda}(n)\mathbf{V}(n)^H \tag{30}$$

where $\mathbf{V}(n) = [\mathbf{v}_1(n)\ \mathbf{v}_2(n)\ \cdots\ \mathbf{v}_{KL}(n)]$ is the $KL \times KL$ orthonormal matrix of eigenvectors, and $\mathbf{\Lambda}(n) = \operatorname{diag}(\lambda_1(n), \lambda_2(n), \cdots, \lambda_{KL}(n))$ is the diagonal matrix of eigenvalues. The r dominant eigenvalues are those eigenvalues that exceed a certain predetermined tolerance level related to the numerical accuracy of the computing device we are using and to the additive noise power. We use the notation $\mathbf{V}(n) = [\mathbf{V}_r(n)\ \mathbf{V}_{KL-r}(n)]$, $\mathbf{\Lambda}(n) = \begin{bmatrix} \mathbf{\Lambda}_r(n) & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{KL-r}(n) \end{bmatrix}$, so that we can define the generalized inverse

$$\left(\tilde{\mathbf{Y}}^{(n)^H}\tilde{\mathbf{Y}}^{(n)}\right)^{\dagger} = \mathbf{V}_r(n)\mathbf{\Lambda}_r(n)^{-1}\mathbf{V}_r(n)^H \tag{31}$$

and the low-rank solution of (29)

$$\tilde{\mathbf{w}}_i(n) = \mathbf{V}_r(n)\mathbf{\Lambda}_r(n)^{-1}\mathbf{V}_r(n)^H\tilde{\mathbf{Y}}^{(n)^H}\tilde{\mathbf{Z}}_i^{(n)}. \tag{32}$$

The *low-rank in-space vector* and *residual vector* are defined, respectively, as

$$\mathbf{Z}_i^{(n)} = \tilde{\mathbf{Y}}^{(n)}\tilde{\mathbf{w}}_i(n) \tag{33}$$

$$\mathbf{u}_i(n) = \tilde{\mathbf{Z}}_i^{(n)} - \mathbf{Z}_i^{(n)}. \tag{34}$$

A. Numerical Procedure to Adaptively Obtain (32)

It is possible to show [24], [7] that if we define $\tilde{\mathbf{Y}}^{(n)^H}\tilde{\mathbf{Y}}^{(n)} = \tilde{\mathbf{\Phi}}(n)$, then an iterative procedure to find an *approximate* eigendecomposition of $\tilde{\mathbf{\Phi}}(n)$ (considered as a time-invariant matrix) is:

- $\mathbf{A}(l) = \tilde{\mathbf{\Phi}}(n)\mathbf{Q}(l-1)$;
- $\mathbf{A}(l) = \mathbf{Q}(l)\tilde{\mathbf{R}}(l)$: square-root decomposition ($\mathbf{Q}(l)$ is constituted by orthonormal columns; $\tilde{\mathbf{R}}(l)$ is upper triangular).

that is

$$\lim_{l\to\infty} \mathbf{Q}(l)\tilde{\mathbf{R}}(l)\mathbf{Q}(l-1)^H \simeq \mathbf{V}_r(n)\mathbf{\Lambda}_r(n)\mathbf{V}_r(n)^H.$$

This is known as simultaneous orthogonal iteration [23]. Strobach [24] and Owsley [16] observed that this iteration can be applied also when $\tilde{\mathbf{\Phi}}(n)$ is slowly varying. The recursive estimation of $\tilde{\mathbf{\Phi}}(n)$, $\mathbf{D}_i(n)$ is

$$\tilde{\mathbf{\Phi}}(n) = \lambda\tilde{\mathbf{\Phi}}(n-1) + (1-\lambda)\tilde{\mathbf{y}}^*(n)\tilde{\mathbf{y}}(n)^T \tag{35}$$

$$\mathbf{D}_i(n) = \lambda\mathbf{D}_i(n-1) + (1-\lambda)\tilde{\mathbf{y}}^*(n)\tilde{z}_i(n) \tag{36}$$

where λ is a parameter close to one that accounts for some exponential weighting [7]. In addition, using the simultaneous orthogonal iteration method and substituting the iteration index l with the time-step index n, it is possible to approximate (31) as

$$\tilde{\mathbf{\Phi}}(n)^{\dagger} = \mathbf{Q}(n-1)\tilde{\mathbf{R}}(n)^{-1}\mathbf{Q}(n)^H. \tag{37}$$

The estimate of the filter weights at time step n is so obtained as

$$\begin{aligned}\tilde{\mathbf{w}}_i(n) &= \mathbf{Q}(n-1)\tilde{\mathbf{R}}(n)^{-1}\mathbf{Q}(n)^H\tilde{\mathbf{Y}}^{(n)^H}\tilde{\mathbf{Z}}_i^{(n)} \\ &= \mathbf{Q}(n-1)\tilde{\mathbf{R}}(n)^{-1}\mathbf{Q}(n)^H\mathbf{D}_i(n).\end{aligned} \tag{38}$$

It was shown in [24] that the update on matrix $\mathbf{A}(n)$ can be performed as

$$\begin{aligned}\mathbf{A}(n) = \lambda\mathbf{A}(n-1)&\mathbf{Q}(n-2)^H\mathbf{Q}(n-1) \\ &+ (1-\lambda)\tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}(n)^H\mathbf{Q}(n-1).\end{aligned} \tag{39}$$

The in-space component, obtained *pinning* the in-space vector, is

$$z_i(n) = \tilde{\mathbf{y}}(n)^T\mathbf{Q}(n-1)\tilde{\mathbf{R}}(n)^{-1}\mathbf{Q}(n)^H\mathbf{D}_i(n) \tag{40}$$

while the residual component, obtained *pinning* the residual vector, is

$$u_i(n) = z_i(n) - \tilde{z}_i(n).$$

Expressions (38)–(40) are at the basis of LORAF1 of [24] as modified for complex signals and applied to our blind deconvolution problem. The algorithm is reported in Table I. The QR decomposition required at every step can be efficiently performed using Jacobi rotations [7]. Observe that $\tilde{\mathbf{w}}_i(n) = \mathbf{Q}(n-1)\mathbf{b}(n)$ is not explicitly computed. Some important refinements of this basic algorithm can be made to improve computational efficiency. Particularly, the explicit updating of the QR decomposition of $\mathbf{A}(n)$ can be eliminated in a scheme which directly tracks the Q and the R factors as new data snapshots are received. This method requires $\frac{r(r+1)(2r+1)}{6}$ Givens rotations per update and is formalized in the algorithm LORAF2 (see [24] for details). Further simplifications can be used to reduce the number of Givens rotations to $2r-1$ and obtain the ultrafast subspace tracking scheme defined by LORAF3 in [24]. The algorithm modified for our problem is reported in Table II. The matrix $\mathbf{G}(n)$ is a Givens rotation matrix [7] that sweeps the last row of the matrix it premultiplies. The structure of the problem (29) may even suggest the use of a square-root-type algorithm for

TABLE I
LORAF1–HOS ALGORITHM

INPUT	
--	$\tilde{z}_i(n) = \left(\lvert z_i(n-1)\rvert^2 - 2\,\sigma_x^2\right)\ z_i(n-1),\ \tilde{\mathbf{y}}(n),\ r,\ \lambda$
SUBSPACE TRACKING	
Step 1	$\mathbf{A}(n) = \lambda\mathbf{A}(n-1)\mathbf{Q}(n-2)^H\mathbf{Q}(n-1) + (1-\lambda)\tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}(n)^H\mathbf{Q}(n-1)$
Step 2	$\mathbf{A}(n) \Rightarrow \mathbf{Q}(n)\tilde{\mathbf{R}}(n)$: QR decomposition
ADAPTIVE FILTERING	
Step 1	$\mathbf{D}_i(n) = \lambda\mathbf{D}_i(n-1) + (1-\lambda)\tilde{\mathbf{y}}^*(n)\tilde{z}_i(n)$
Step 2	Solve by back substitution $\tilde{\mathbf{R}}(n)\mathbf{b}(n) = \mathbf{Q}(n)^H\mathbf{D}_i(n)$ for $\mathbf{b}(n)$
Step 3	$z_i(n) = \mathbf{b}(n)^T\mathbf{Q}(n-1)^H\tilde{\mathbf{y}}(n)$
INITIALIZATION	
Step 1	$\mathbf{Q}(n-2) = \mathbf{Q}(n-1) = \begin{bmatrix} \mathbf{I}_r \\ 0 \end{bmatrix}$
Step 2	$\mathbf{A}(n-1) = \mathbf{Q}(n-2)\tilde{\mathbf{y}}(n-1)\tilde{\mathbf{y}}(n-1)^H$
Step 3	$\mathbf{D}_i(n-1) = \tilde{\mathbf{y}}^*(n-1)\tilde{z}_i(n-1)$

subspace tracking—this approach would have the immediate advantage of reducing the required dynamic range of the triangular factor $\tilde{\mathbf{R}}(n)$.

B. Adaptive Dominant Space Order Estimation

The order r of the dominant space can be estimated using a threshold level tol which, added to the estimated noise floor (of the additive white Gaussian noise), will help to discriminate the relevant eigenvectors. The subspace section of the algorithm is operated with the order r equal to $r_{\max}$. This maximum subspace order $r_{\max}$ is experimentally determined so that the rank of the autocovariance matrix statistically never exceeds this parameter. A possible adaptive noise power estimator is

$$\sigma_n^2(n) = KLp_y(n) - \sigma_x^2$$

with

$$p_y(n) = \lambda p_y(n-1) + \frac{1-\lambda}{KL}\,\mathrm{tr}[\tilde{\mathbf{y}}(n)\tilde{\mathbf{y}}(n)^H].$$

Only the eigenvectors corresponding to the eigenvalues that exceed $\sigma_n^2(n) + \mathrm{tol}$ are used in the computation of the filter $\tilde{\mathbf{w}}_i(n)$ and as a consequence in signal reconstruction. Observe that the eigenvalues (actually their approximation) are obtained by extracting the diagonal elements of the upper triangular matrix $\tilde{\mathbf{R}}(n)$ (which is an *almost diagonal* matrix). The experimental results of the next section do not make use, however, of the adaptive space order estimation.[8] The order $r = r_{\max}$ is actually experimentally determined offline for a certain propagation scenario as a tradeoff between error rate performance and computational complexity.

C. Remarks on Implementation

The simulations in the following section are performed using a word length size of 24 bits, using fixed-point arithmetic.

[8]There is marginal performance improvement if real-time estimation and tracking of r is performed in the fixed number of mobile transmitters scenario, which is the case that we simulated. Adaptive order estimation becomes indispensable when the number of transmitters U varies dynamically—the rank properties of the space–time autocorrelation matrix may change dramatically in this case.

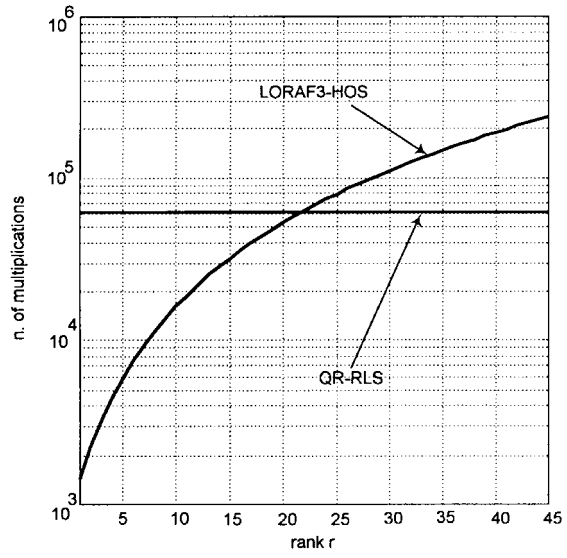
TABLE II
LORAF3–HOS ALGORITHM. THE MATRIX $\mathbf{G}(n)$ IS A GIVENS ROTATION MATRIX

INPUT	
--	$\tilde{z}_i(n) = \left(\lvert z_i(n-1)\rvert^2 - 2\,\sigma_x^2\right)\ z_i(n-1),\ \tilde{\mathbf{y}}(n),\ r,\ \lambda$
SUBSPACE TRACKING	
Step 1	$\mathbf{h}(n) = \mathbf{Q}(n-1)^H\mathbf{y}(n)$
Step 2	$\mathbf{y}_\perp(n) = \tilde{\mathbf{y}}(n) - \mathbf{Q}(n-1)\mathbf{h}(n)$
Step 3	$\overline{\mathbf{y}}_\perp(n) = \frac{\mathbf{y}_\perp(n)}{\lVert\mathbf{y}_\perp(n)\rVert}$
Step 4	$\begin{bmatrix} \tilde{\mathbf{R}}(n) \\ 0^T \end{bmatrix} \Leftarrow \mathbf{G}(n) \begin{bmatrix} \lambda\tilde{\mathbf{R}}(n-1) + (1-\lambda)\mathbf{h}(n)\mathbf{h}(n)^H \\ (1-\lambda)\mathbf{h}(n)^H\lVert\mathbf{y}_\perp(n)\rVert \end{bmatrix}$
Step 5	$\begin{bmatrix} \mathbf{Q}(n) & \mathbf{q}(n) \end{bmatrix} \Leftarrow \begin{bmatrix} \mathbf{Q}(n-1) & \overline{\mathbf{y}}_\perp(n) \end{bmatrix} \mathbf{G}(n)^H$
ADAPTIVE FILTERING	
Step 1	$\mathbf{D}_i(n) = \lambda\mathbf{D}_i(n-1) + (1-\lambda)\tilde{\mathbf{y}}^*(n)\tilde{z}_i(n)$
Step 2	$\mathbf{f}_i(n) = \lambda\mathbf{f}_i^+(n-1) + (1-\lambda)\mathbf{h}(n)\tilde{z}_i(n)$
Step 3	$\begin{bmatrix} \mathbf{f}_i^+(n) \\ \times \end{bmatrix} = \mathbf{G}(n) \begin{bmatrix} \mathbf{f}_i(n) \\ \overline{\mathbf{y}}_\perp(n)^H\mathbf{D}_i(n) \end{bmatrix}$
Step 4	Solve by back substitution $\tilde{\mathbf{R}}(n)\mathbf{b}(n) = \mathbf{f}_i^+(n)$ for $\mathbf{b}(n)$
Step 5	$z_i(n) = \mathbf{b}(n)^T\mathbf{h}(n)$
INITIALIZATION	
Step 1	$\mathbf{Q}(n-1) = \begin{bmatrix} \mathbf{I}_r \\ 0 \end{bmatrix}$
Step 2	$\mathbf{f}_i^+(n-1) = 0$
Step 3	$\tilde{\mathbf{R}}(n-1) = 0$
Step 4	$\mathbf{D}_i(n-1) = \tilde{\mathbf{y}}^*(n-1)\tilde{z}_i(n-1)$

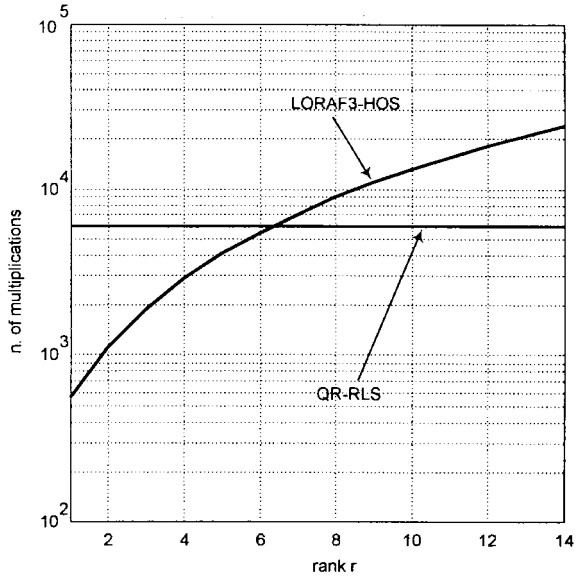
TABLE III
COMPUTATIONAL COMPLEXITY ANALYSIS WITH K ELEMENTS, L TAPS, AND RANK APPROXIMATION EQUAL TO r

Method	LORAF3-HOS	QR-RLS
N. of real multiplications	$8KLr + 106r^2 + 124r + 20KL - 58$	$10KL + 30(KL)^2$
N. of Givens Rotations	$2r - 1$	KL

The computational complexity in terms of multiplications and elementary complex 2×2 Givens rotation per update was calculated and compared to the complexity of the adaptive QR-RLS [7]. The approximate number of computations per iteration using filters length equal to L and K sensors is given in Table III.



(a)



(b)

Fig. 1. The LORAF1-HOS algorithm. Computational complexity per update of LORAF3-HOS and QR-RLS in terms of multiplications and versus the subspace order retained in the adaptive algorithm. (a) Computational complexity versus rank reduction (a) $L = 9, K = 5$ and (b) $L = 9, K = 2$.

It is evident that the LORAF3-HOS algorithm here proposed becomes convenient in terms of computational complexity with respect to the QR-RLS scheme if the order r of the dominant space remains below a certain value—in Fig. 1 for $K = 5$ and $K = 2$ and $L = 9$, the number of multiplications required by the two algorithms is computed and plotted versus the order r. LORAF3-HOS is convenient for $r \leq 21$ for $K = 5$ and for $r \leq 6$ for $K = 2$.

V. Results of Experiments

A TDMA system for cellular communications has been simulated according to [26] and [27]. In addition, we present the results of some lab experiments performed using the Watkins–Johnson wide-band dual-mode (AMPS and IS-136)

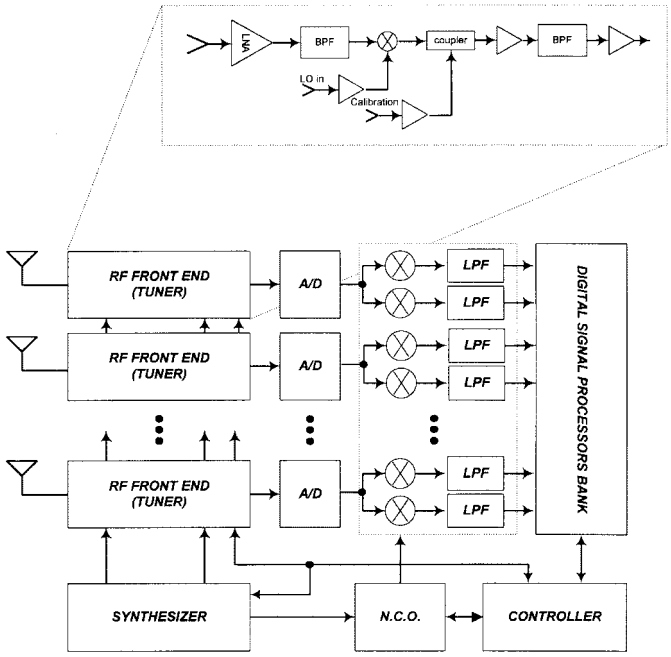


Fig. 2. Block diagram of the receiver.

base-station system Base_2. The main purpose of experimenting the algorithm with data samples collected from hardware equipment is to demonstrate practical applicability of the ideas presented. A block diagram of the receiver section of the base station is shown in Fig. 2. The tuner module performs a standard single conversion scheme. The analog-to-digital (A/D) is a high-speed bandpass sampler, while the conversion at baseband is operated by digital downconverters (wide-band processing). The slots are 162 symbols long, and 14 symbols at the beginning of each slot are known at the receiver.[9] It is obvious that no blind algorithm can achieve convergence in 14 symbols. The scheme we propose is a hybrid scheme which during training performs low-rank least-squares filtering, while after convergence runs in blind mode as described previously. The difference is only in the desired signal that is fed back to the adaptive algorithm. During training the desired signal is $x_i^{\mathrm{train}}(n)$ instead of $\tilde{z}_i(n)$, where $x_i^{\mathrm{train}}(n)$ is the training sequence, that is

$$\tilde{z}_i(n) = \begin{cases} x_i^{\mathrm{train}}(n), & \text{training} \\ \left(|z_i(n-1)|^2 - 2\sigma_x^2\right) z_i(n-1), & \text{blind.} \end{cases} \tag{41}$$

The mean-squared error (MSE) is defined as the average of the squared error obtained over $\tilde{M} = 100$ Monte Carlo runs and is given by $\mathrm{MSE_i} = \frac{1}{\tilde{\mathrm{M}}} \sum_{\nu=1}^{\tilde{\mathrm{M}}} |\epsilon_{\mathrm{i}}^{(\nu)}(\mathrm{n})|^2$. The error obtained at the mth run is $\epsilon_i^{(\nu)}(n) = z_i^{(\nu)}(n) - x_i(n - n_i)$, where n_i is the delay introduced by the filters and $z_i^{(\nu)}(n)$ is the output of the combined filters relative to the ith transmitter obtained at the νth run. Observe that the initial training with known symbols associated with the user ID matching procedure solves the sources permutation ambiguity and the phase uncertainty. Of particular importance is the fact

[9] This corresponds to a frame format similar to the DTC of [26] and [27].

TABLE IV
CHANNEL PROPAGATION ENVIRONMENTS FOR PERFORMANCE EVALUATION RESULTS: DOA'S AND DELAY SPREADS OF THE MOBILES. IN ALL CASES $\tau_{1,1}^{(1)} = \tau_{2,1}^{(1)} = \tau_{1,1}^{(2)} = \tau_{2,1}^{(2)} = \tau_{1,1}^{(3)} = \tau_{2,1}^{(3)} = 0$

	K, L, r	U	θ_2, θ_3 ($\theta_1 = 0°$)	$\tau_{1,2}^{(1)}$	$\tau_{2,2}^{(1)}$	$\tau_{1,2}^{(2)}$	$\tau_{2,2}^{(2)}$	$\tau_{1,2}^{(3)}$	$\tau_{2,2}^{(3)}$
Fig. 4	5, 9, 22	3	30°, −70°	41.2 μs	41.2 μs	10.3 μs	10.3 μs	5.15 μs	5.15 μs
Fig. 5 (a)	2, 3, 4	2	45°, −90°	10.3 μs	10.3 μs	10.3 μs	10.3 μs	10.3 μs	10.3 μs
Fig. 5 (b)	2, 3, 4	2	45°, −90°	20.6 μs	0.0 μs	10.3 μs	10.3 μs	10.3 μs	10.3 μs
Fig. 5 (c)	2, 3, 4	2	30°, −70°	41.2 μs	0.0 μs	0.0 μs	10.3 μs	10.3 μs	10.3 μs
Fig. 5 (d)	2, 3, 4	2	30°, −70°	41.2 μs	41.2 μs	10.3 μs	10.3 μs	10.3 μs	10.3 μs
Fig. 5 (e)	2, 3, 4	2	30°, −70°	41.2 μs	41.2 μs	41.2 μs	41.2 μs	41.2 μs	41.2 μs
Fig. 6 (a)	2, 5, 6	1	− − . − −	41.2 μs	0.0 μs	− − . − −	− − . − −	− − . − −	− − . − −
Fig. 6 (b)	2, 5, 6	1	− − . − −	41.2 μs	0.0 μs	− − . − −	− − . − −	− − . − −	− − . − −
Fig. 7 (a)	2, 5, 6	1	− − . − −	41.2 μs	0.0 μs	− − . − −	− − . − −	− − . − −	− − . − −
Fig. 7 (b)	2, 5, 6	1	− − . − −	41.2 μs	0.0 μs	− − . − −	− − . − −	− − . − −	− − . − −
Fig. 8	5, 3 : 7, 15	3	10°, −45°	41.2 μs	0.0 μs	41.2 μs	0.0 μs	41.2 μs	0.0 μs
Fig. 10	2, 9, 12	1	− − . − −	41.2 μs	0.0 μs	− − . − −	− − . − −	− − . − −	− − . − −

TABLE V
CHANNEL PROPAGATION ENVIRONMENTS FOR PERFORMANCE EVALUATION RESULTS: VELOCITY OF THE MOBILES

	V_1	V_2	V_3	λ_{QR-RLS}	$\lambda_{LORAF3-HOS}$
Fig. 4	8 Km/hr	8 Km/hr	8 Km/hr	− − . − −	0.97
Fig. 5 (a)	50 Km/hr	50 Km/hr	8 Km/hr	− − . − −	0.9
Fig. 5 (b)	50 Km/hr	50 Km/hr	8 Km/hr	− − . − −	0.9
Fig. 5 (c)	100 Km/hr	100 Km/hr	100 Km/hr	− − . − −	0.9
Fig. 5 (d)	75 Km/hr	50 Km/hr	8 Km/hr	− − . − −	0.9
Fig. 5 (e)	100 Km/hr	100 Km/hr	8 Km/hr	− − . − −	0.87
Fig. 6 (a)	100 Km/hr	− − . − −	− − . − −	− − . − −	0.9
Fig. 6 (b)	100 Km/hr	− − . − −	− − . − −	− − . − −	0.9
Fig. 7 (a)	100 Km/hr	8 Km/hr	8 Km/hr	0.855	0.87
Fig. 7 (b)	50/100 Km/hr	8 Km/hr	− − . − −	0.855	0.87
Fig. 8	100 Km/hr	50 Km/hr	50 Km/hr	0.855/0.9	0.855/0.9
Fig. 10	8/50/100 Km/hr	− − . − −	− − . − −	0.87	0.855

that we always compare the proposed approach LORAF3-HOS with a more traditional QR-RLS approach, explicitly a second-order-statistics-based method. The QR-RLS uses the synchronization sequences while in training and past decisions while in decision-directed mode.

A. Computer Simulations Results

In the simulations we assumed a sensor spacing of $\tilde{\lambda}/2$. The number of elements and transmitters for each figure is specified in Table IV. One of the transmitters is at array broadside and is the signal of interest. We assume a two-path model, and each path impulse response is modeled as a two-ray Rayleigh-fading channel ($P_1 = P_2 = 2$) [26], [27]. The arrival angles of the paths are spread around 0° with a cluster width of 2°. The interfering signals are generated with the same parameters but with DOA's clustered around θ_2 and θ_3, as indicated in Table IV. In Fig. 3 the equivalent baseband discrete-time model is shown relative to the ith mobile transmitter. Delay spread propagation parameters are summarized in Table IV as they relate to the test cases reported in the figures. Observe that the symbol period is 41.2 μs and that in the described model $\tau_{1,1}^{(1)} = \tau_{2,1}^{(1)} = \tau_{1,1}^{(2)} = \tau_{2,1}^{(2)} = \tau_{1,1}^{(3)} = \tau_{2,1}^{(3)} = 0$ μs. The Doppler frequency usually describes the second-order statistics of channel variations. Doppler frequency is related through wavelength $\tilde{\lambda}$ to the ith mobile transmitter velocity V_i expressed in km/h. The model used in this case is based on the wide sense stationary uncorrelated scattering (WSSUS) assumption [3]. The complex weights 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.[10] Table V summarizes the Doppler frequency situation as related to the test cases' results shown in the figures. The SNR

[10] The Doppler spectrum is approximated 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 Doppler spectrum.

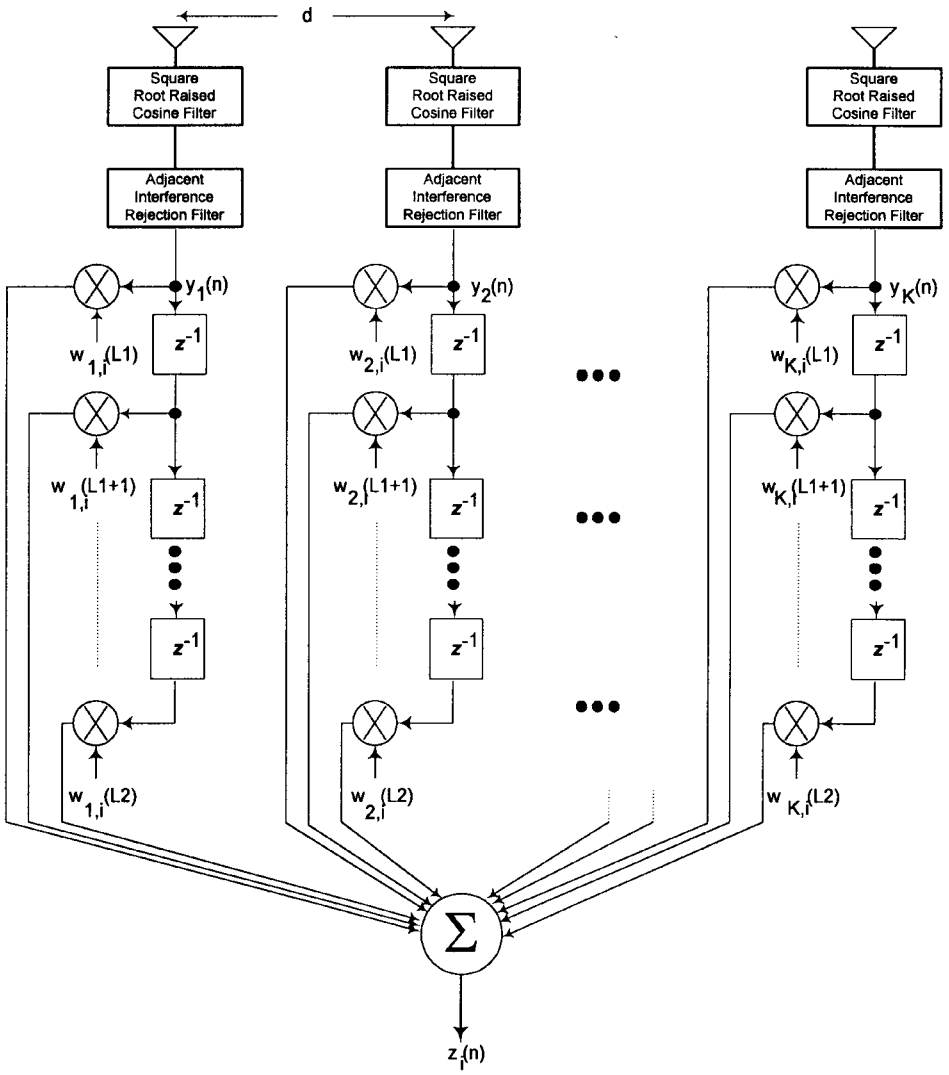


Fig. 3. Discrete-time model of the filtering section (K sensors) relative to the ith transmitter. Observe that an adjacent interference-rejection filter is concatenated with the (square-root) raised-cosine filter, which by itself does not meet the IS-136 specification in terms of out-of-band rejection.

for each discrete-time channel impulse response is defined as in [2]. Fig. 4 shows an experiment where the filters are running on experimental data (not formatted as in the IS-136 TDMA system) in blind adaption using LORAF1-HOS mode for the entire sequence. The speed is relatively low but large delay spread is present on the channel, while the cochannel interference suppression capability of the method is evident (the three signals are received with the same average power). It is, however, also important to note that the experiment shows the weakness of the algorithm—slow startup convergence. Unfortunately, the filter reaches satisfactory MSE only after hundreds of symbols. The solution to the slow startup problem is, as already mentioned, training, according to the scheme (41). Fig. 5 shows experiments (again with no relation to a realistic TDMA frame) where we try to investigate the tracking capability of the method. The plots show the trajectory of the real part of the center tap[11] as compared to the center tap of the ideal distortionless reception filter (constrained to have finite length), computed assuming perfect instantaneous knowledge of $\mathcal{H}(z)$ and shown as the dashed trace. The procedure we used to compute the ideal optimum setting for $\tilde{\mathbf{w}}_i(n)$ was the following:

1) *freeze* time evolution of the multipath channels parameters at time step nT;
2) compute the discrete-time symbol-spaced channels $h_{l,k}(m)$ for $m = J_1, J_1+1 \cdots J_2$ using (3) and compact into matrix $\tilde{\mathbf{H}}$ as in (13), say $\tilde{\mathbf{H}}(n)$;
3) solve

$$\min_{\tilde{\mathbf{w}}_i(n)} ||\tilde{\mathbf{H}}(n)\tilde{\mathbf{w}}_i(n) - \tilde{\boldsymbol{\delta}}_i||^2$$

using

$$\tilde{\mathbf{w}}_i(n) = (\tilde{\mathbf{H}}(n)^H \tilde{\mathbf{H}}(n))^{\dagger} \tilde{\mathbf{H}}(n)^H \tilde{\boldsymbol{\delta}}_i.$$

Fig. 5(a)–(d) shows results for infinite SNR and low delay spread of the interfering signals (see Table IV). Fig. 5(d) shows the traces of the real part of two adjacent taps. Fig. 5(e) is for root-mean-square (rms) delay spread of 20.6 μs for all signals impinging over the array. Fig. 6 shows ideal experiments with one single transmitter and periodic retraining with randomly

[11] In equalization literature the *center tap* is the tap that carries the maximum energy.

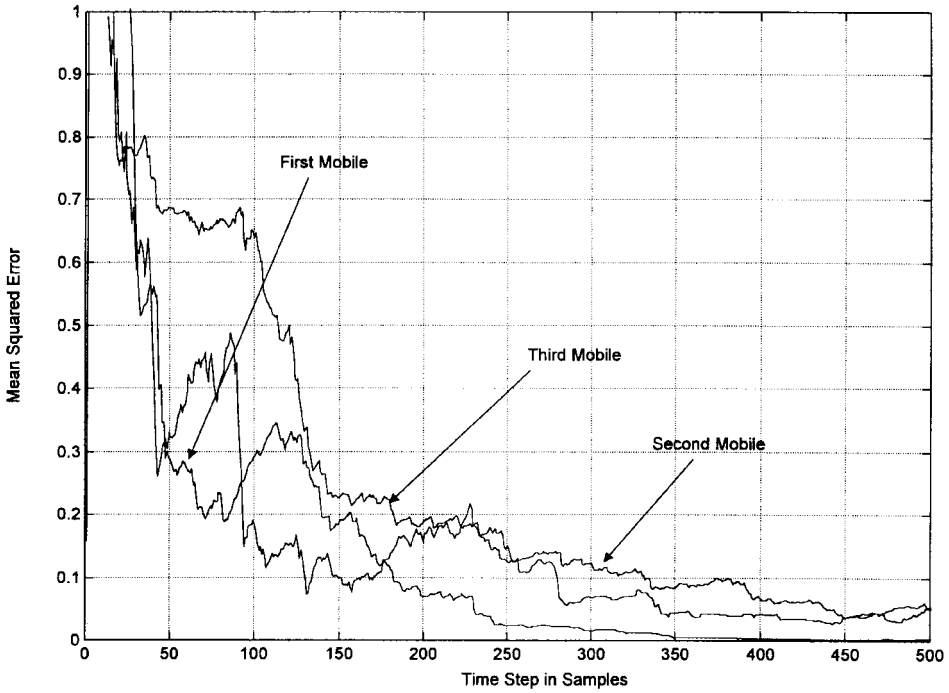


Fig. 4. MSE versus time step for the three transmitters with blind adaption (using LORAF1-HOS) during the entire length of the transmitted sequence. For propagation parameters, see Tables IV and V.

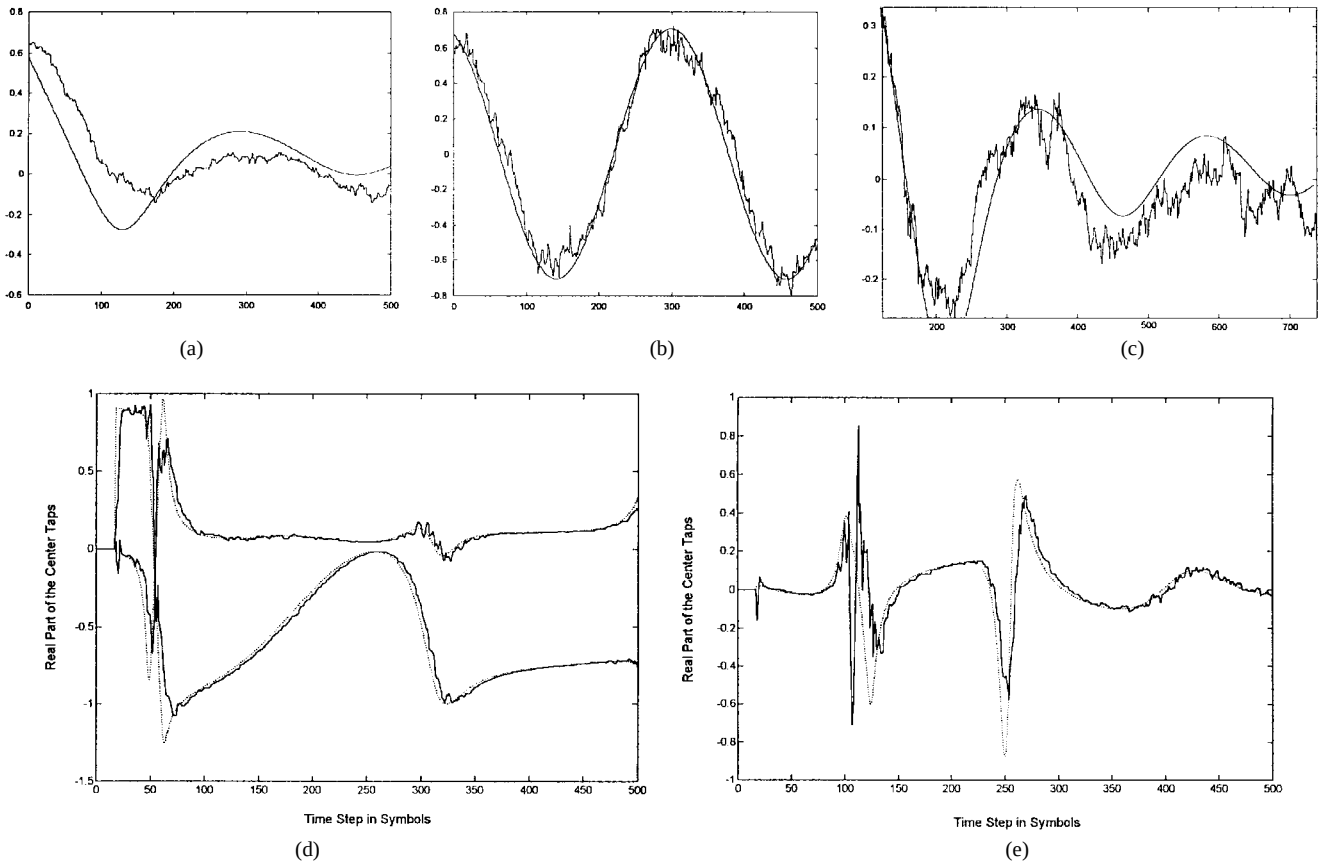


Fig. 5. An experiment that shows the tracking capability of the algorithm LORAF3-HOS. The smooth dashed trace is the optimum Wiener filter solution, computed assuming perfect knowledge of the discrete-time channel impulse response. For propagation parameters, see Tables IV and V. SNR is infinite.

generated inputs (not IS-136 frames). Each case depicts: 1) the Rayleigh-fading amplitude evolution in time of the two-ray Rayleigh-fading channel; 2) the real part of the center tap and the theoretical behavior of the ideal distortionless reception filter; and 3) the magnitude of the error $|\epsilon(n)|$ for the single transmitter. Several important comments are in order. Observe

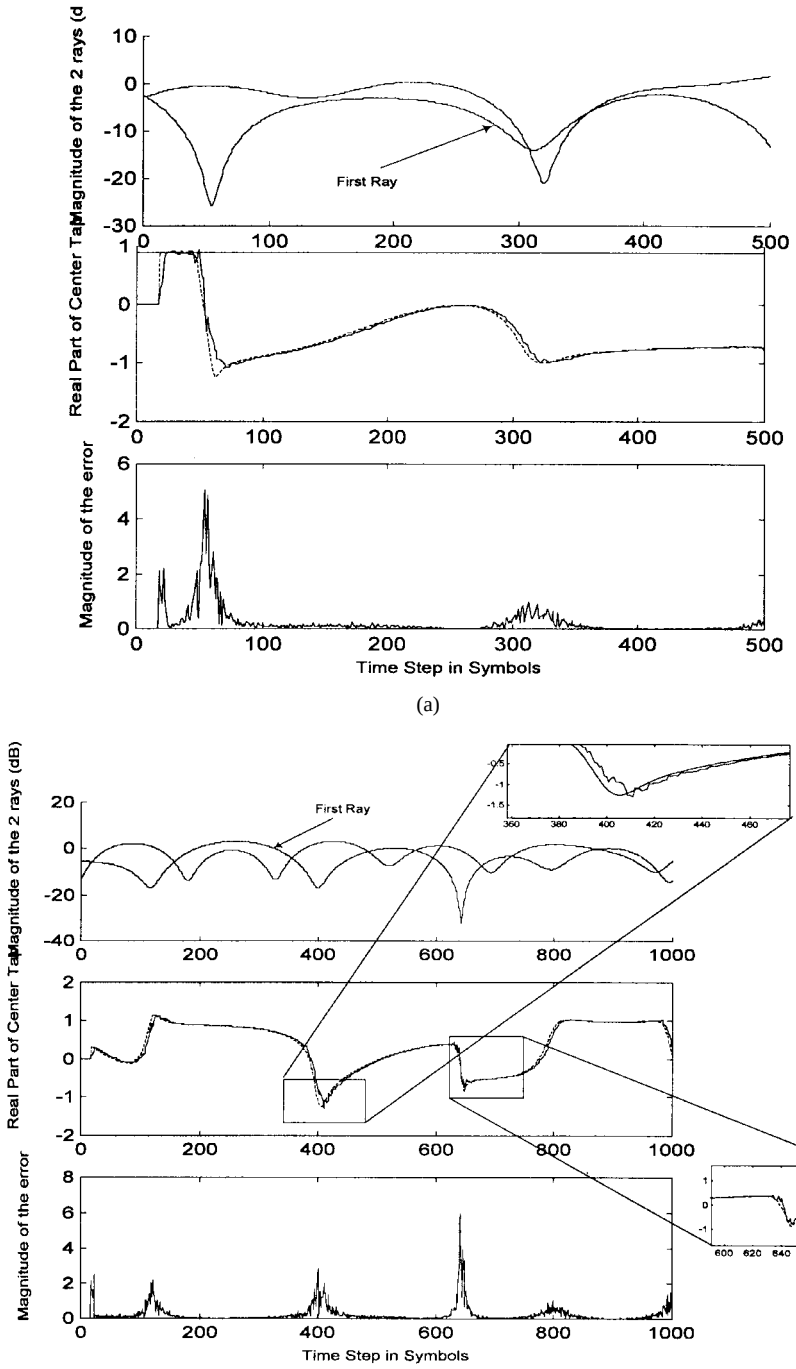
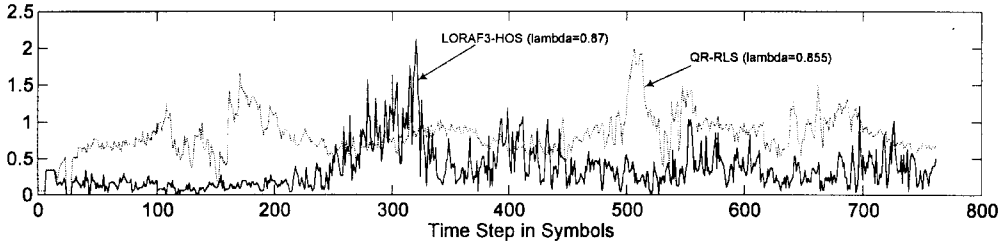
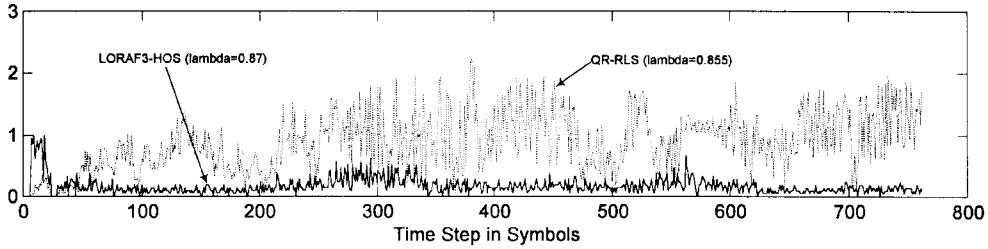


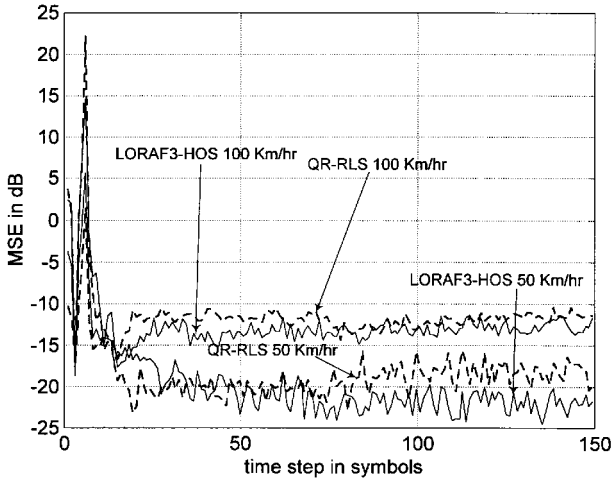
Fig. 6. For one single transmitter, plots of the two-ray Rayleigh-fading channel amplitudes, real part of the center tap of the filter for LORAF3-HOS versus the optimum Wiener solution (dashed line), and magnitude of the error. The filter is periodically retrained as the MSE reaches a threshold. Evidently, fading events of the first ray dramatically affect performance. For propagation parameters, see Tables IV and V.

that fades of the first ray below −10 dB dramatically impact the performance of the adaptive algorithm. In some works the concept of traditional equalizers (based on second-order statistics) not able to identify nonminimum phase channels was emphasized. Our results indicate that the use of HOS cannot totally solve the problem because of the fast variations of the parameters. The algorithm is extremely sensitive to fades occurring on the first ray, or if the amplitude of the second ray becomes larger than the first one. This problem can be explained by the fact that as the first ray decreases in amplitude, the center tap should move to an adjacent tap, which is a configuration of the taps considerably different from the original one (observe, in fact, the rapid changes in the real part of the center tap in Fig. 6 in correspondence with a fade of the first ray). The filter has to deconvolve a channel whose spectral response has dramatically changed and it should be *repositioned in time* with respect to the maximum energy point captured by the synchronizer. Since the synchronization point is not changed dynamically within one slot, the filter is not able to satisfactorily deconvolve the channel. The degradation





(a)



(b)

Fig. 7. Stability experiments (a) for the two algorithms at $\mathrm{SNR} = 21$ dB: QR-RLS and LORAF3-HOS at 100 km/h. The slot is 750 symbols (for testing purposes only). Top: center coefficient after synchronization. Bottom: instantaneous magnitude of the error after synchronization. (b) The averaged MSE (100 Monte Carlo runs) is shown for real demodulation of IS-136 slots. For propagation parameters, see Tables IV and V.

could be probably reduced, increasing the time span of the filter (larger L), or by introducing a feedback filtering section. We reserve the investigation of this important topic to future work. Fig. 7(b) shows the MSE averaged over 100 independent computer runs versus time step in symbols for $\frac{E_b}{N_0} = 21$ dB for real IS-136 slots, while Fig. 7(a) is an experiment with a 750–symbol-long sequence and *no retraining*. It is shown how the blind mode is more robust under fast fading than decision-directed mode of traditional schemes such as the QR-RLS algorithm. Bit-error rate (BER) analysis results are shown in Fig. 8 (at different speeds) with delays as specified in Table IV. BER is relative to the first mobile transmitter. Ideal frame and symbol synchronization is assumed. The SNR is the same on each discrete-time channel. The two cochannel interferers are received at an average power 5 dB below the average power of the first transmitter. A sample size of 10^{k+4} was used to estimate an error probability of 10^{-k}. Note that for $r = 15$, $L = 3$, the computational complexity of LORAF3-HOS is approximately four times the QR-RLS complexity; for $r = 15$, $L = 5$, it is about 1.5 times the QR-RLS complexity; and for $r = 15$, $L = 7$, it is about 20% less computationally intensive. While we obtain improvement in all cases, the remarkable result is that for $L = 7$, LORAF3-HOS gives for $\mathrm{BER} = 10^{-2}$ an SNR-per-bit improvement of about 6 dB.

B. Hardware Implementation Results

A simpler and indeed more realistic scenario ($K = 2$, $U = 1$)[12] was studied using data collected from the digital signal processing (DSP) receiver section of Base_2, the dual-mode wide-band base station implemented at Watkins-Johnson Company. The hardware test setup is depicted in Fig. 9; the propagation parameters are in Table IV. A hardware multipath fading simulator is connected to the two antenna ports of the base station. The IS-136 signal generator simulates transmission of digital traffic channel (DTC) frames coming

[12] This testing environment is also specified in [26] and [27].

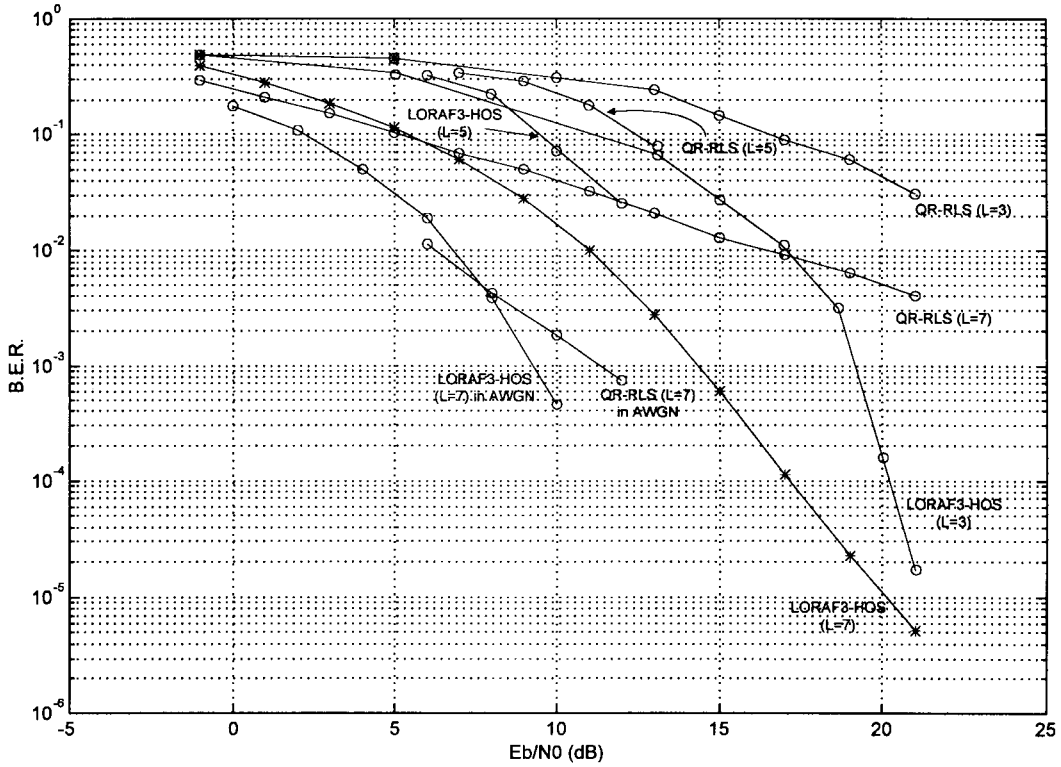


Fig. 8. BER of the first mobile transmitter in a three-mobile environment. For propagation parameters, see Tables IV and V. Also, results for the no-fading case are shown.

from three different mobiles. Additive Gaussian noise is injected on both diversity channels. Observe that the DSP modem receives a sampling rate of 80 kHz (not an integer of the symbol rate). The wordlength used is 24 and the algorithm has been implemented using simulated fixed-point arithmetic.[13] A polyphase raised-cosine filter concatenated with an adjacent interference rejection filter transforms the rate to $4/T = 97.2$ kHz. Then the two-channel filter works at $2/T$ rate. AGC is operated on a slot-by-slot basis. While perfect synchronization was assumed in the previous simulations, in the hardware experiments there is an open-loop synchronizer[14] choosing the optimum positioning of the filters at $4/T$ rate. In addition, as required by the specification [26], [27], there is carrier frequency offset between the local oscillator and the transmitted carrier frequency of about 213 Hz. The adaptive filter is not able to track this large frequency offset. Indeed, frequency offset is more conveniently estimated using a second-order phase-locked loop (PLL) that compensates for the frequency drift. A detailed description of the synchronizer and the frequency offset compensator is omitted because it is beyond the scope of this paper. The results of extensive BER measurements are summarized in Fig. 10. The improvement with respect to the QR-RLS is not as impressive as in the previous experiments, where multiple interference sources and a relatively large number of sensors made tracking the optimum MMSE solution more difficult. It is exactly in these circumstances that LORAF3-HOS outperforms known solutions in terms of performance and computational complexity. It is important to mention, however, that substantial improvement in the two-antenna case can be achieved using QR-based decision-feedback schemes and similar extensions of the cumulant-based algorithm.

VI. Conclusion

We have studied a new practical solution to the array processing problem in a cellular base station employing antenna arrays. The method is based on a vector generalization [13] of the idea presented in [21] and the low-rank adaptive processing concept of [24]. The algorithm is sufficiently fast to track channel variations caused by moving transmitters, while at the same time being highly attractive from the computational point of view, proving that the use of HOS does not necessarily imply slow convergence and, hence, extremely large sample size. Fast startup convergence is achieved by the use of training sequences. The experimental investigation is of particular interest because the applicability of the new method was verified in a realistic environment specified by the current U.S. standard for digital cellular communications [26], [27]. BER results were shown using also data collected from hardware.

The algorithm LORAF3-HOS is perfectly adequate to track channel variations due to Doppler shifts larger than 80 Hz more efficiently than traditional decision-directed-based methods, especially in multiple cochannel interference and fast-fading

[13] A simulation analysis of the dynamic range required not to degrade the performance of the algorithm as opposed to the floating point implementation was carried out, although the description of such analysis is beyond the scope of this paper. Different variables of the algorithm required different number of fractional bits to avoid overflow. In particular, the matrix $\bar{\mathbf{R}}(n)$ required double-precision representation.

[14] The basic strategy is to cross correlate the incoming signal with the synchronization sequences adequately interpolated.

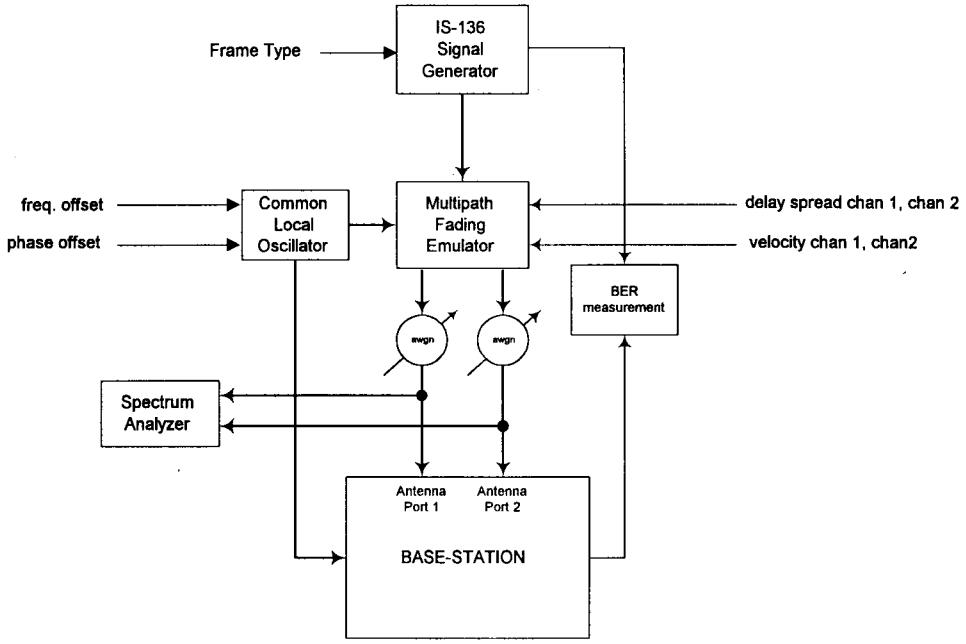


Fig. 9. Hardware test setup for laboratory experiments.

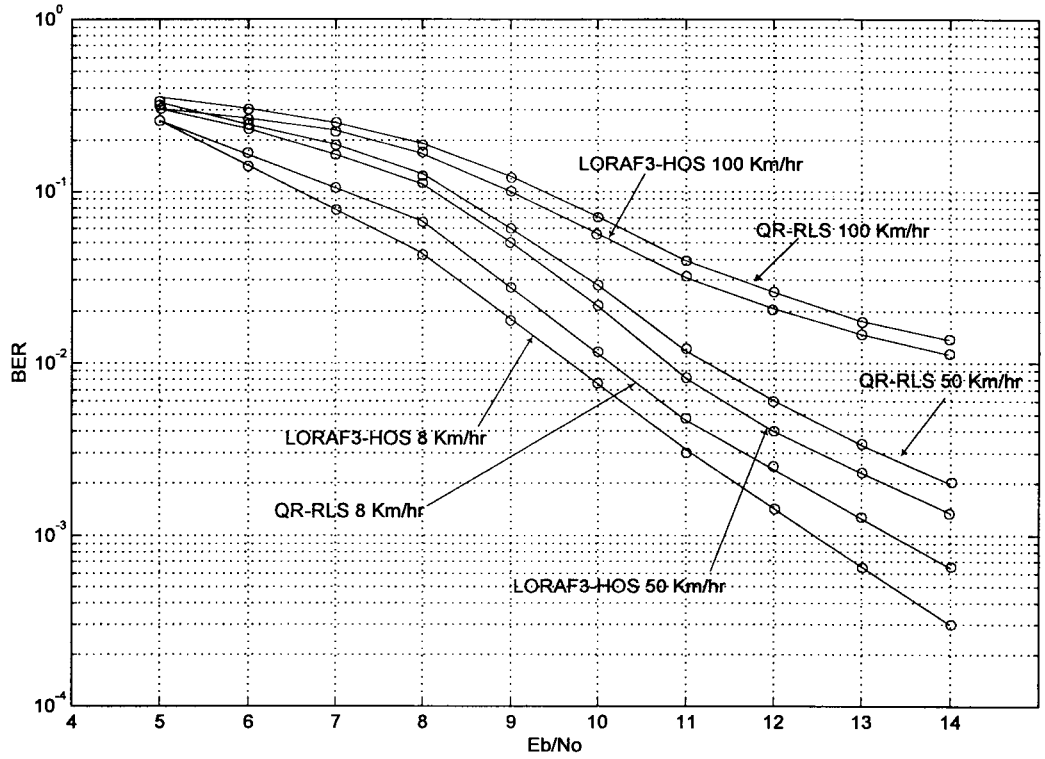


Fig. 10. BER for the hardware experiments. For propagation parameters, see Tables IV and V.

scenarios when using a relatively large number of sensors. In these cases, tracking the optimum MMSE solution using traditional methods such as RLS or its variations is not only computationally intensive but also not satisfactory in terms of performance, due to the rank degeneracy problem of the space–time correlation matrix. The best known method for RLS filtering in terms of numerical properties, the QR-RLS was compared to the proposed algorithm, LORAF3-HOS. The remarkable result is that at $\mathrm{BER} = 10^{-2}$, the improvement of LORAF3-HOS in SNR per bit with respect to the traditional decision-directed QR-RLS is about 6 dB, using same length of the filters and same number of sensors. Moreover, only 80% of the computational complexity per update of the full-rank QR-RLS is required to obtain such improvement. The computational complexity reduction that one can obtain with the low-rank tracking idea is very attractive; in fact, it is pos-

sible to obtain satisfactory BER results (compliant with [26] and [27]) using less than 60% of the QR-RLS computational complexity.

APPENDIX
ON THE CONVERGENCE OF THE ALGORITHM

The iterative procedure (15) and (16) applied to the vector $\tilde{\mathbf{s}}_i$ maintains the index of the tap with largest magnitude, which was called in [21] the *leading tap* (see for details [21, Sec. III]). 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 nonzero 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 transmitter and not an arbitrary j_0th transmitter. Although this problem in our algorithm has been solved by use of training sequences, the convergence to the ith transmitter can be ensured 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 $\mathbf{w}_{i,j} = \mathbf{0}$ 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 problem of the convergence of a certain function $\mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i) = \mathbf{\Psi}(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i)$ (when $\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i$) to the same solution of the function $\mathbf{\Psi}(\tilde{\mathbf{s}}_i)$ is a well-known and investigated problem in the blind deconvolution literature (see [22, chapter by Shalvi and Weinstein]). This convergence cannot always be guaranteed for both functions. First of all, observe that the procedure (15), (16) is a gradient-based search to solve the maximization problem

$$\max_{\tilde{\mathbf{s}}_i} \mathbf{\Psi}(\tilde{\mathbf{s}}_i) \tag{42}$$

subject to the constraint $||\tilde{\mathbf{s}}_i||^2 = 1$. In fact, the two steps in (15) and (16) are equivalent to the gradient-based iteration

$$\tilde{\mathbf{s}}_i = \tilde{\mathbf{s}}_i + \mu \frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i} \tag{43}$$

with $\mathbf{\Psi}(\tilde{\mathbf{s}}_i) = \sum_j \sum_k s_{i,j}(k)^2 s_{i,j}(k)^{*2}$. $\frac{\partial f(\mathbf{v})}{\partial \mathbf{v}}$ indicates the gradient of the vector $\mathbf{v}$ and we choose a *very large* step size μ. However, we have translated the maximization procedure for $\mathbf{\Psi}(\tilde{\mathbf{s}}_i)$ over $\tilde{\mathbf{s}}_i$ into a maximization for a certain $\mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i)$ over $\tilde{\mathbf{w}}_i$, where

$$\mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i) = \mathbf{\Psi}(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i) = \mathbf{\Psi}(\tilde{\mathbf{s}}_i)$$

and we have used $\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i$.

Let us assume that $\tilde{\mathbf{s}}_i^E$ is an extremum for $\mathbf{\Psi}(\tilde{\mathbf{s}}_i)$, that is, $\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}|_{\tilde{\mathbf{s}}_i = \tilde{\mathbf{s}}_i^E} = \mathbf{0}$. It is then obviously true that $\tilde{\mathbf{w}}_i^E$ such that $\tilde{\mathbf{s}}_i^E = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i^E$ is also an extremum for $\mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i)$ because $\frac{\partial \mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i)}{\partial \tilde{\mathbf{w}}_i}|_{\tilde{\mathbf{w}}_i = \tilde{\mathbf{w}}_i^E} = \mathbf{0}$.

The converse may not be true. That is, if we assume that $\bar{\tilde{\mathbf{w}}}_i^E$ is an extremum for $\mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i)$, it may not be true that $\bar{\tilde{\mathbf{s}}}_i^E = \tilde{\mathbf{H}}\bar{\tilde{\mathbf{w}}}_i^E$ is an extremum for $\mathbf{\Psi}(\tilde{\mathbf{s}}_i)$. In fact, we may have $\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i)}{\partial \tilde{\mathbf{w}}_i}|_{\tilde{\mathbf{w}}_i = \bar{\tilde{\mathbf{w}}}_i^E} = \mathbf{0}$ if $\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}|_{\tilde{\mathbf{s}}_i^E = \tilde{\mathbf{H}}\bar{\tilde{\mathbf{w}}}_i^E} \neq \mathbf{0}$ belongs to the kernel of $\tilde{\mathbf{H}}$ which is orthogonal to the subspace spanned by $\tilde{\mathbf{H}}$, which means that $\bar{\tilde{\mathbf{w}}}_i^E$ can be far from the desired solution.

We investigate this last important issue. By the chain rule we have

$$\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i)}{\partial w_{i,n}(m)} = \sum_{k=1}^{U} \sum_{l} h_{n,k}^*(m-l) \cdot \frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial s_{i,k}(l)}\bigg|_{s_{i,k}(l) = \sum_{k_1=1}^{K} \sum_{m_1} w_{i,k_1}(m_1) h_{k_1,k}(l-m_1)} \tag{44}$$

or, in vector form

$$\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i)}{\partial \tilde{\mathbf{w}}_i} = \tilde{\mathbf{H}}^H \frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}\bigg|_{\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i}. \tag{45}$$

Now multiply both sides of (44) by $\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^\dagger$ to obtain

$$\begin{aligned} \tilde{\mathbf{H}}(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^\dagger \frac{\partial \mathbf{\Psi}(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i)}{\partial \tilde{\mathbf{w}}_i} &\simeq \mathbf{I}_U \frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}\bigg|_{\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i} \\ &= \frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}\bigg|_{\tilde{\mathbf{s}}_i = \tilde{\mathbf{H}}\tilde{\mathbf{w}}_i} \end{aligned} \tag{46}$$

because $\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^\dagger\tilde{\mathbf{H}}^H \simeq \mathbf{I}_U$, the $U \times U$ identity matrix. It is then evident from (46) that if $\bar{\tilde{\mathbf{w}}}_i^E$ is an extremum for $\mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i)$, that is, if $\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{H}}\tilde{\mathbf{w}}_i)}{\partial \tilde{\mathbf{w}}_i}|_{\tilde{\mathbf{w}}_i = \bar{\tilde{\mathbf{w}}}_i^E} = \mathbf{0}$, then $\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}|_{\tilde{\mathbf{s}}_i^E = \tilde{\mathbf{H}}\bar{\tilde{\mathbf{w}}}_i^E} \simeq \mathbf{0}$. Since $||\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}||$ is continuous with respect to $w_{i,n}(m)$ and $h_{k,n}(m)$, then we only need to require that for a sufficiently small $\Delta > 0$ it is satisfied

$$||\tilde{\mathbf{H}}(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^\dagger\tilde{\mathbf{H}}^H - \mathbf{I}_U||_F < \Delta \tag{47}$$

to guarantee that $||\frac{\partial \mathbf{\Psi}(\tilde{\mathbf{s}}_i)}{\partial \tilde{\mathbf{s}}_i}||$ is arbitrarily small, which implies that there exists an extremum $\tilde{\mathbf{s}}_i^E$ such that $||\bar{\tilde{\mathbf{s}}}_i^E - \tilde{\mathbf{s}}_i^E||$ is arbitrarily small. In other words, the extremum for $\mathbf{\Psi}_{\tilde{\mathbf{w}}_i}(\tilde{\mathbf{w}}_i)$ is arbitrarily close to the extremum for $\mathbf{\Psi}(\tilde{\mathbf{s}}_i)$ if there exists a sufficiently small Δ such that (47) is verified. Condition (47) emphasizes the importance of selecting the rank order of the low-rank filter $r = r_{\max}$ carefully. The selection of a low value for r will alleviate the computational effort of the algorithm but will make the algorithm prone to misconvergence, due to (47).

ACKNOWLEDGMENT

The author would like to thank the Base_2 technical staff of the Telecommunications Group, Watkins-Johnson Company, for some enlightening discussions.

REFERENCES

[1] S. Anderson, M. Millnert, M. Viberg, and B. Wahlberg, "An adaptive array for mobile communication systems," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 230–236, Feb. 1991. .

[2] P. Balaban and J. Salz, "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.

[3] P. A. Bello, "Characterization of randomly time variant linear channels," *IEEE Trans. Commun.*, vol. COM-11, pp. 360–393, Dec. 1963.

[4] 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. AC-23, pp. 1043–1047, Dec. 1978.

[5] J. M. Cioffi and T. Kailath, "Fast recursive least squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 302–337, Apr. 1984.
[6] Z. Ding, "On convergence analysis of fractionally spaced adaptive blind equalizers," *IEEE Trans. Signal Processing*, vol. 45, pp. 650–657, Mar. 1997.
[7] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
[8] D. Hatzinakos and C. L. Nikias, "Blind equalization using a tricepstrum based algorithm," *IEEE Trans. Commun.*, vol. 39, pp. 669–682, Aug. 1991.
[9] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.
[10] Y. Li and K. J. Ray Liu, "Blind MIMO FIR channel identification based on second-order statistics with multiple signal recovery," Univ. Maryland, College Park, Tech. Rep. ISR TR 96-7, Nov. 1995; *IEEE Trans. Inform. Theory*, submitted for publication.
[11] ———, "On blind equalization of MIMO channels," *ICC'96 Proceedings*, Dallas, TX, 1996, pp. 1000–1004.
[12] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 43, pp. 2120–2129, Sept. 1995.
[13] M. Martone, "Non-Gaussian multivariate adaptive AR estimation using the super exponential algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 2640–2644, Oct. 1996.
[14] 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.
[15] 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.
[16] N. L. Owsley, "Adaptive data orthogonalization," in *Proc. IEEE ICASSP*, Tampa, FL, 1978, pp. 109–112.
[17] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[18] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989.
[19] M. Rosenblatt, *Stationary Sequences and Random Fields*. Boston, MA: Birkhauser, 1985.
[20] L. L. Scharf, *Statistical Signal Processing*. New York: Addison-Wesley, 1991.
[21] O. Shalvi and E. Weinstein, "Super exponential methods for blind deconvolution," *IEEE Trans. Inform. Theory*, vol. 39, pp. 504–519, Mar. 1993.
[22] ———, "Universal methods for blind deconvolution," in *Blind Deconvolution*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall 1994.
[23] G. W. Stewart, "Methods of simultaneous iteration for calculating eigenvectors of matrices," in *Topics in Numerical Analysis II*, J. H. Miller, Ed. New York: Academic, 1975, pp. 169–185.
[24] P. Strobach, "Low-rank adaptive filters," *IEEE Trans. Signal Processing*, vol. 44, pp. 2932–2947, Dec. 1996.
[25] A. Swami, G. Giannakis, and S. Shamshunder, "Multichannel ARMA processes," *IEEE Trans. Signal Processing*, vol. 42, pp. 898–913, Apr. 1994.
[26] *TDMA Cellular/PCS—Radio Interface—Mobile Station—Base Station Compatibility—Digital Control Channel*, TIA/EIA/IS-136.1-A, Oct. 1996.
[27] *TDMA Cellular/PCS—Radio Interface—Mobile Station—Base Station Compatibility—Traffic Channels and FSK Control Channel*, TIA/EIA/IS-136.2-A, Oct. 1996.
[28] 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.
[29] J. H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 2, pp. 579–586, June 1987.

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

From 1990 to 1991 he was with the Italian Air Force and also 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 Technical Staff of the On-Board Equipment Division, Alenia Spazio, where he was involved in the design of satellite receivers and spread spectrum telemetry tracking and control transponders. He also collaborated with the Digital Communications Research Group, Fondazione Ugo Bordoni. In 1994 he was appointed Visiting Scientist at the Electrical, Computer, Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY. He was a Wireless Communications Consultant for ATS, Inc., Waltham, MA, and in 1995 he joined the Telecommunications Group, Watkins-Johnson Company, Gaithersburg, MD. He is also with the Electrical Engineering Department, George Washington University, Washington, DC. His main interests are in advanced signal processing for wireless receivers implementation, spread-spectrum multiple-access communications, and cellular radio architectures.

Dr. Martone is a member of the New York Academy of Sciences and the American Association for the Advancement of Science.