On MMSE Real-Time Antenna Array Processing Using Fourth-Order Statistics in the U.S. Cellular TDMA System

Massimiliano (Max) Martone, *Member, IEEE*

***Abstract*—The antenna array processing problem in the reverse link of the current U.S. digital cellular communication system is studied and higher-than-second-order-statistics (HOS) baseband processing is proposed as a possible candidate solution. The remarkable difference of our approach as compared to other existing similar techniques is the idea of the minimization of the mean squared error using fourth-order cumulants alone and *nonblind* criteria. A recursive Jacobi total least squares algorithm is used in the adaptive implementation to mitigate the effects of high error variance in the estimates of the cumulants based on sample statistics. The method is shown to be very effective in a fast fading environment with multiple cochannel interferers.**

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

I. Introduction

IN WIDEBAND time division multiple access (TDMA) systems, data dispersion can span several symbols as a consequence of frequency selective fading caused by RF multipath propagation. The received signal is composed of the original plus several delayed replicas, and each replica reaches the antenna with different attenuation and angle of arrival. In addition, multiple cochannel interferers may be received at the antenna afflicted by similar impairments. Since propagation characteristics of the signal of interest and interference change in time due to the motion of the transmitters, it is of paramount importance to employ adequate real-time signal processing at the base-station receiver of the cellular system in order to maintain the highest signal quality. Space-only processing methods [23] are not effective in such an environment because intersymbol interference (ISI) cannot be compensated using the traditional combining architecture. Space-time filtering may result in a more efficient approach to mitigate multipath fading and interference caused by multiple cochannel transmitters. On the other hand, as a consequence of the continuous decrease in the cost of digital hardware, there is today great interest in those techniques that process digitized baseband samples coming from different sensors of an antenna array. One of the possible solutions to the problem is the application of the minimum mean squared error (MMSE) principle [2], which usually results in multichannel generalizations of single-channel adaptive linear filtering algorithms. Unfortunately, the MMSE method has limitations, despite its simplicity, because second-order statistics (SOS) alone are processed which imply the assumption that the underlying processes are all Gaussian processes. Since in a real-world scenario this assumption is rarely satisfied, the use of SOS results in suboptimum performance. The use of higher-than-second-order statistics (HOS) has generated great interest in the signal processing community over the last ten years [8], [17], [19], mainly because of the celebrated theoretical Gaussian rejection property and enhanced identification capability. Most parts of the works describing HOS-based algorithms for identification/deconvolution are blind and try not to exploit any input signal knowledge. Many successful applications of HOS were recently proposed in a multichannel setup [14]–[16]. However, since the estimation of HOS requires considerably larger sample size than SOS, these blind algorithms result in an intolerably slow convergence behavior that prevented their applicability in practical systems. In TDMA systems (see, for example, the IS-136 standard for U.S. cellular communications [21]), the bitstream (control data and voice data) is organized into frames. Some known sequences are periodically transmitted in every slot that allows the receiver to perform frame synchronization, symbol timing recovery, and equalizer training. The length of a slot including the training sequence is usually in the range 100–200 symbols, depending on the type of frame, while the length of the known sequence is in the range 14–20 symbols. The use of the known sequence is indispensable to fast start-up equalization. The method proposed in this work is based on the same idea introduced in [5] and [6] where high-order cumulants were used to derive MMSE methods for system identification. Here we modify the basic intuition so that it can be applied to the deconvolution of complex vector signals. Moreover, we show that it is indeed possible to derive a fully adaptive implementation based on a recursive Jacobi-type algorithm. 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 and laboratory hardware experiments are shown for IS-136 [21], the current digital standard for cellular communications in the United States.

Manuscript received September 1, 1997; revised February 1, 1998.

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

Publisher Item Identifier S 0733-8716(98)07891-3.

II. System Model

A brief review of the system model is given in this section. We assume U mobile transmitters communicating

with a base-station with a K-element antenna array, with $U \leq K$. Each element of the antenna has a digital filter with $L = L_2 - L_1 + 1$ complex weights. The structure of the antenna is assumed to be a uniform linear array, d is the distance between adjacent antenna elements, $\tilde{\lambda}$ is the wavelength of the signal. Multipath propagation for the lth mobile transmitter can be characterized as an $N^{(l)}$-ray channel whose nth ray ($n = 1, 2, \cdots, N^{(l)}$) is represented by $P_n^{(l)}$ received delayed and attenuated replicas of the signal. The impulse response of the nth ray relative to the lth transmitter can be expressed as $f_n^{(l)}(t) = \sum_{m=1}^{P_n^{(l)}} \rho_{n,m}^{(l)} e^{j\psi_{n,m}^{(l)}} \delta(t - \tau_{n,m}^{(l)})$, where $\tau_{n,m}^{(l)}$, $\rho_{n,m}^{(l)}$, and $\psi_{n,m}^{(l)}$ are delay, amplitude, and phase of the mth delayed signal in the nth path relative to the lth transmitter, while $\delta(t)$ is the delta function.[1] Observe that we are assuming for the derivation a time invariant channel, while instead usually $\tau_{n,m}^{(l)}$, $\rho_{n,m}^{(l)}$, and $\psi_{n,m}^{(l)}$ 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 $m_l(t) = \sum_k a_l(k) p_{tx}(t - kT)$, where $a_l(k) = \tilde{a}_k^{(l)} + j\tilde{b}_k^{(l)}$ 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 roll-off factor 0.35, and T is the signaling interval. The symbols are assumed to have predetermined correlation properties. The model assumed for the generation of the symbols is $a_l(n) = \sum_m b_l(m) x_l(n - m)$, where $b_l(m)$ is a finite impulse response (FIR) filter modeling the autocorrelation of the complex sequence $a_l(k)$. The statistical properties of $x_l(n)$ are detailed later. The lth transmitted signal propagated through the nth path can be represented as $S_n^{(l)}(t) = \int_{-\infty}^{+\infty} f_n^{(l)}(t - \tau) m_l(\tau)\, d\tau = \sum_{m=1}^{P_n^{(l)}} \rho_{n,m}^{(l)} e^{j\psi_{n,m}^{(l)}} m(t - \tau_{n,m}^{(l)}) e^{j2\pi f_0 (t - \tau_{n,m}^{(l)})}$ where $\omega_0 = 2\pi f_0$ is the carrier frequency. The contribution of the lth transmitted signal propagated through the nth path with angle of arrival (DOA) $\theta_n^{(l)}$ and phase difference $e^{-j2\pi k\, d(\sin\theta_n^{(l)}/\tilde{\lambda})}$ from the first antenna element to the kth element can be written (we are neglecting the additive noise) as $r_k(t, \theta_n^{(l)}) = e^{j\omega_0 t} \sum_{m=1}^{P_n^{(l)}} \rho_{n,m}^{(l)} m_l(t - \tau_{n,m}) e^{-j2\pi k\, d(\sin\theta_n^{(l)}/\tilde{\lambda})} e^{j\phi_{n,m}^{(l)}}$ where $\phi_{n,m}^{(l)} = -2\pi f_0 \tau_{n,m}^{(l)} + \psi_{n,m}^{(l)}$. 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

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

where $\eta_k(n)$ is Gaussian noise and $h_{k,l}(m) = h_{k,l}(mT)$ is the T-sampled[3] impulse response $h_{k,l}(t) = \sum_{n=1}^{N_l} \sum_{m=1}^{P_n^{(l)}} \rho_{n,m}^{(l)} r_p(t - \tau_{n,m}^{(l)}) e^{j\phi_{n,m}^{(l)}} e^{-j2\pi k\, d(\sin\theta_n^{(l)}/\tilde{\lambda})}$. In this expression, $r_p(t)$ is the raised cosine function with excess bandwidth 0.35 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. Discrete-time convolution is indicated as $*$. 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. To indicate statistics estimated in a noise-free environment we will use notation $\langle\cdots\rangle_{\mathrm{SNR}=\infty}$. For example, if $r(t) = s(t) + n(t)$ where $n(t)$ is noise and $s(t)$ is signal of interest, we have $\langle E\{r(t) r^*(t + \tau)\}\rangle_{\mathrm{SNR}=\infty} = E\{s(t) s^*(t + \tau)\}$.

In the z-domain, the transfer function (1) can be expressed as $\tilde{\mathcal{H}}(z) = \sum_{k=-\infty}^{\infty} \mathbf{H}(k) z^{-k}$ where the organization of the $\mathcal{H}_{i,l}(z)$ polynomials (z-transforms of $h_{i,l}(k)$) 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 $\tilde{\mathcal{W}}(z) = \sum_{k=L_1}^{L_2} \mathbf{W}(k) z^{-k}$ with length $L = L_2 - L_1 + 1$ is applied to the output of the array. The main 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

$$\tilde{\mathcal{W}}(z) \tilde{\mathcal{H}}(z) = \mathbf{I}_U \tag{2}$$

where $\mathbf{I}_U$ is a $U \times U$ identity matrix. The system $\tilde{\mathcal{W}}(z)$ is required to be bounded-input bounded-output (BIBO) stable. The solution (2) 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 reconstruction is possible only up to a constant delay, due to the stationarity of the input process*. The recovered signals will

[1] In this model the κth ray 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 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 [21] we have $\tilde{a}_m^{(l)} = \tilde{a}_{m-1}^{(l)} \cos[\Delta\phi_m^{(l)}] - \tilde{b}_{m-1}^{(l)} \sin[\Delta\phi_m^{(l)}]$, $\tilde{b}_m^{(l)} = \tilde{a}_{m-1}^{(l)} \sin[\Delta\phi_m^{(l)}] + \tilde{b}_{m-1}^{(l)} \cos[\Delta\phi_m^{(l)}]$, where $\Delta\phi_m^{(l)} = \pi/4$ if $bit_{1,m}^{(l)} = 0$ and $bit_{2,m}^{(l)} = 0$, $\Delta\phi_m^{(l)} = 3\pi/4$ if $bit_{1,m}^{(l)} = 1$ and $bit_{2,m}^{(l)} = 0$, $\Delta\phi_m = -3\pi/4$ if $bit_{1,m} = 1$ and $bit_{2,m} = 1$, $\Delta\phi_m = -\pi/4$ if $bit_{1,m} = 0$ and $bit_{2,m} = 1$.

[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.

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

$$\tilde{\mathcal{W}}(z)\tilde{\mathcal{H}}(z) = \mathbf{P}\mathcal{D}(z) \tag{3}$$

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

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

with $\phi_i \in [-\pi, \pi]$, n_i integer for $i = 1, 2, \cdots, U$. We say that $\tilde{\mathcal{H}}(z)$ satisfies the *distortionless reception* condition if for $\tilde{\mathcal{H}}(z)$ it 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 [12], [13]

$$\det(\tilde{\mathcal{H}}^H(e^{j\omega})\tilde{\mathcal{H}}(e^{j\omega})) \neq 0, \qquad \text{for all } \omega \in [-\pi, \pi]. \tag{4}$$

In the time domain, the linear filter can be written

$$z_i(n) = \sum_{l=1}^{K}\sum_{m=L_1}^{L_2} w_{i,l}(m)y_l(n-m) \qquad i = 1, 2, \cdots, U \tag{5}$$

where $w_{i,l}(m)$ is the filter corresponding to the polynomial $\mathcal{W}_{i,l}(z)$. Neglecting the additive noise terms, the overall impulse response is characterized by the input/output relation

$$z_i(n) = \sum_{l=1}^{U}\sum_{m} s_{i,l}(m)a_l(n-m) \qquad i = 1, 2, \cdots, U \tag{6}$$

where $a_l(n) = \sum_m b_l(m)x_l(n-m)$ and $s_{i,l_1}(m_2) = \sum_{l=1}^{K}\sum_{m=L_1}^{L_2} w_{i,l}(m)h_{l,l_1}(m_2-m)$.

The *desired* response $\tilde{\mathbf{s}}_i = [\mathbf{s}_{i,1}^T, \mathbf{s}_{i,2}^T, \mathbf{s}_{i,3}^T, \cdots, \mathbf{s}_{i,U}^T]^T$, $\mathbf{s}_{i,l} = [\cdots, s_{i,l}(-1), s_{i,l}(0), s_{i,l}(1), \cdots]^T$, that completely restores the information signal of the ith transmitter up to the delay n_i, can be expressed as

$$\tilde{\boldsymbol{\delta}}_i = [\boldsymbol{\delta}_{i,1}^T, \boldsymbol{\delta}_{i,2}^T, \boldsymbol{\delta}_{i,3}^T, \cdots, \boldsymbol{\delta}_{i,U}^T]^T,$$
$$\boldsymbol{\delta}_{i,j} = \begin{cases} (\cdots, 0, 0, 0, \cdots), & i \neq j \\ \hat{\boldsymbol{\delta}}_i, & i = j. \end{cases} \tag{7}$$

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 ($\mathbf{P} = \mathbf{I}_U$). Our task is to design a deconvolution filter with taps $w_{i,l}(m) l = 1, 2, \cdots, K$, $m = L_1, L_1+1, \cdots, L_2$ such that $\tilde{\mathbf{s}}_i$ approximates in some sense $\tilde{\boldsymbol{\delta}}_i$.

B. Key Assumptions

The fundamental assumptions necessary to develop the algorithm are as follows.

- **AS1:** The transformation in (1) represents a stable system.
- **AS2:** $\tilde{\mathcal{H}}(z)$ is *irreducible* and $\mathbf{H}(k) \neq \mathbf{0}$ only for $k \in [J_1, J_2]$ $(J = J_2 - J_1 + 1)$ with $\mathbf{H}(J_2)$ full rank.
- **AS3:** The complex sequences $\{a_i(n)\}$ are mutually independent and are all generated by linear stable systems

$$a_i(n) = b_i(n) * x_i(n) \qquad i = 1, 2, \cdots, U$$

with finite impulse responses $b_i(k)$ $(M_1 \leq k \leq M_2)$ of length $M = M_2 - M_1 + 1$ with $b_i(M_1) \neq 0$, $b_i(M_2) \neq 0$ for any $i = 1, 2, \cdots, U$ and no zero on the unit circle. The cumulants of $\{x_i(n)\}$, independently identically distributed (i.i.d.) non-Gaussian processes, satisfy

—$E\{x_i(n)\} = E\{x_i^3(n)\} = 0$,
—$\text{cum}[x_i(n), x_l^*(n)] = \sigma_x^2 > 0$, only for $i = l$.
—$E\{x_i(n)x_l(n)\} = E\{x_i^*(n)x_l^*(n)\} = 0$, for any i, l,
—$\text{cum}[x_{l_1}(n), x_{l_2}(n), x_{l_3}^*(n), x_{l_4}^*(n)] = \gamma_{4x} \neq 0$, only for $l_1 = l_2 = l_3 = l_4$.

Assumption **AS2** is required to assure the *distortionless reception* condition for $\tilde{\mathcal{H}}(z)$. We observe, in fact, that **AS2** is equivalent to stating that the *generalized Sylvester matrix* $H_{L,J}$ of the MIMO system [4], [11] is full column rank (proof of this fact is in [12], [13]). This implies that a *distortionless reception* linear filter exists (i.e., $\tilde{\mathcal{H}}(z)$ satisfies the *distortionless reception* condition). Observe that this assumption appears reasonable in the channel model assumption described in [21]. Assumption **AS3** generalizes the algorithm to handle colored input processes [10] and attempts to model the fact that the source signals in practice do not satisfy the usual i.i.d. assumption. This is true not only in training but also in decision-directed mode. For example, in [21], 14 symbols at the beginning of any slot are dedicated to synchronization and eventual training. These symbols constitute complex *sounding* sequences which satisfy particular autocorrelation properties. They only approximate the correlation shape of optimal *sounding* sequences. Moreover, in decision-directed mode, symbols coming from the digital traffic channel [21] do not satisfy the ideal i.i.d. assumption. It is, however, extremely important to emphasize that:

1) the algorithm we propose does not require the impulse responses $b_i(n)$ to be known;
2) it applies with no change (only marginal modifications in the derivation are necessary) also to the case when $a_l(n)$ can be considered i.i.d. processes, that is when $a_l(n) = x_l(n)$.

III. Derivation of the Algorithm

The restoration error for the ith channel is defined as

$$\epsilon_i(n) = a_i(n) - \sum_{l=1}^{K}\sum_{m} w_{i,l}(m)y_l(n-m)$$
$$= a_i(n) - \sum_{l=1}^{U} s_{i,l}(n) * a_l(n) \tag{8}$$

where we have neglected the contribution of the additive noise. The usual approach is to find the weights $w_{i,l}(m)$ for $l = 1, 2, \cdots, K$, $m = L_1, L_1+1, \cdots, L_2$ such that the mean squared error (MSE) is minimized. The noise-free MSE in our

model can be expressed as

$$\begin{aligned}\langle E\{|\epsilon_i(n)|^2\}\rangle_{\mathrm{SNR}=\infty} &= E\left\{\left|a_i(n)-\sum_{l=1}^{U}s_{i,l}(n)*a_l(n)\right|^2\right\}\\ &= E\left\{\left|\sum_{l=1}^{U}\sum_{k}\tilde{s}_{i,l}(k)x_l(n-k)\right|^2\right\}\\ &= \sigma_x^2\sum_{l=1}^{U}\sum_{k}|\tilde{s}_{i,l}(k)|^2 \qquad (9)\end{aligned}$$

where $\tilde{s}_{i,l}(k)=\sum_m[\delta_{i,l}(m)-s_{i,l}(m)]b_l(k-m)$, and we have defined in (7) $\boldsymbol{\delta}_{i,l}=[\cdots,\delta_{i,l}(-1),\delta_{i,l}(0),\delta_{i,l}(1),\cdots]^T$, so that $a_i(n)$ can be written as $a_i(n)=\sum_l\sum_m\delta_{i,l}(m)a_i(n-m)$. The MSE as expressed in (9) shows that it is possible to obtain some cumulant-based criteria for deconvolution whose objective is the minimization of a cost function proportional to (9). If we organize the weights $w_{i,l}(m)$ in a vector

$$\begin{aligned}\tilde{\mathbf{w}}_i &= \left[\mathbf{w}_{i,1}^T,\mathbf{w}_{i,2}^T,\cdots,\mathbf{w}_{i,K}^T\right]^T\\ \mathbf{w}_{i,l} &= [w_{i,l}(L_1),w_{i,l}(L_1+1),\cdots,w_{i,l}(L_2)]^T\end{aligned}$$

the following theorem can be proved.

Theorem 1: The minimization of the cost function

$$J^{(i)}(\tilde{\mathbf{w}}_i)=\left|\sum_{l=1}^{K}\sum_{k=-\infty}^{+\infty}\mathrm{cum}[\epsilon_i(n),\epsilon_i^*(n),y_l(n-k),y_l^*(n-k)]\right|^2$$

with respect to $w_{i,l}(k)$, $l=1,2,\cdots,K$, $k=L_1,\cdots,L_2$ is equivalent to the minimization of the noise-free mean squared error $\langle E\{|\epsilon_i(n)|^2\}\rangle_{\mathrm{SNR}=\infty}$ up to a scalar factor if

$$\sum_{l_1=1}^{K}\sum_{k}|g_{l_1,i_1}(k)|^2=\sum_{l_2=1}^{K}\sum_{k}|g_{l_2,i_2}(k)|^2=\frac{\mathcal{K}}{\gamma_{4x}}\neq 0$$

for any $i_1,i_2\in[1,2,\cdots,U]$ with $g_{i,l}(k)=h_{i,l}(k)*b_l(k)$.

Proof: Using the fact that $\epsilon_i(n)=\sum_{l=1}^{U}\sum_k\tilde{s}_{i,l}(k)x_l(n-k)+$ *noise terms* [see (9)], and the noise-rejection properties of fourth-order cumulants, the cost function $J^{(i)}(\tilde{\mathbf{w}}_i)$ can be expressed as

$$\begin{aligned}J^{(i)}(\tilde{\mathbf{w}}_i) &= \left|\gamma_{4x}\sum_{l_1=1}^{K}\sum_{k}\sum_{l_2=1}^{U}\sum_{m}|\tilde{s}_{i,l_2}(m)|^2|g_{l_1,l_2}(m-k)|^2\right|^2\\ &= \left|\gamma_{4x}\sum_{l_2=1}^{U}\sum_{m}|\tilde{s}_{i,l_2}(m)|^2\sum_{l_1=1}^{K}\sum_{k}|g_{l_1,r}(k)|^2\right|^2\\ &= \left|\sigma_x^2\sum_{l=1}^{U}\sum_{m}|\tilde{s}_{i,l}(m)|^2\right|^2\left|\frac{\mathcal{K}}{\sigma_x^2}\right|^2 \qquad (10)\end{aligned}$$

where $\mathcal{K}=\gamma_{4x}\sum_{l=1}^{K}\sum_k|g_{l,r}(k)|^2$. Now it is evident from (10) as compared to (9) that, as long as $\mathcal{K}$ is different from zero, minimizing $J^{(i)}(\tilde{\mathbf{w}}_i)$ with respect to $w_{i,l}(k)$, $l=1,2,\cdots,K$, $k=L_1,\cdots,L_2$ is equivalent to minimize $\langle E\{|\epsilon_i(n)|^2\}\rangle_{\mathrm{SNR}=\infty}$. Q.E.D.

Observe that the condition $\mathcal{K}/\gamma_{4x}\neq 0$ of the Theorem holds due to **AS2** and **AS3**, while $\sum_{l_1=1}^{K}\sum_k|g_{l_1,i_1}(k)|^2=\sum_{l_2=1}^{K}\sum_k|g_{l_2,i_2}(k)|^2$ for any $i_1,i_2\in[1,2,\cdots,U]$ can be forced in practice using independent automatic gain control (AGC) circuits.

Theorem 2: The optimum deconvolution filter associated with the minimum of the cost function $J^{(i)}(\tilde{\mathbf{w}}_i)$ satisfies

$$\sum_{l_1}\sum_{k}\mathrm{cum}\left[\epsilon_i(n),y_{l_2}^*(n-m),y_{l_1}(n-k),y_{l_1}^*(n-k)\right]=0 \qquad (11)$$

and

$$\sum_{l_1}\sum_{k}\mathrm{cum}\left[\epsilon_i^*(n),y_{l_2}(n-m),y_{l_1}(n-k),y_{l_1}^*(n-k)\right]=0 \qquad (12)$$

for $l_2=1,2,\cdots,K$, $m=L_1,L_1+1,\cdots,L_2$, $i=1,2,\cdots,U$.

The proof of this theorem is reported in Appendix A.

Using (11), it is then true that

$$\begin{aligned}\sum_{l_1}\sum_{k}&\mathrm{cum}\left[a_i(n),y_{l_2}^*(n-m),y_{l_1}(n-k),y_{l_1}^*(n-k)\right]\\ &-\sum_{l_3}\sum_{m_1}w_{i,l_3}(m_1)\sum_{l_1}\sum_{k}\mathrm{cum}\left[y_{l_3}(n-m_1),\right.\\ &\left.\times y_{l_2}^*(n-m),y_{l_1}(n-k),y_{l_1}(n-k)\right]=0\end{aligned}$$

and we can write the following equations:

$$c_{a,y}^{(i_1,l_1)}(m)=\sum_{l_2=1}^{K}\sum_{m_1=L_1}^{L_2}w_{i_1,l_2}(m_1)c_{y,y}^{(l_2,l_1)}(m-m_1) \qquad (13)$$

for $i_1=1,2,\cdots,U$ $l_1=1,2,\cdots,K$ $m=L_1,L_1+1,\cdots,L_2$, where

$$\begin{aligned}&c_{a,y}^{(i_1,i_2)}(m)\\ &=\sum_{l=1}^{K}\sum_{k=D_1^{(m)}}^{D_2^{(m)}}\mathrm{cum}\left[a_{i_1}(n),y_{i_2}^*(n-m),y_l(n-k),y_l^*(n-k)\right]\\ &c_{y,y}^{(i_1,i_2)}(m)\\ &=\sum_{l=1}^{K}\sum_{k=B_1^{(m)}}^{B_2^{(m)}}\mathrm{cum}\left[y_{i_1}(n),y_{i_2}^*(n-m),y_l(n-k)y_l^*(n-k)\right]\end{aligned}$$

where $B_1^{(m)}$, $B_2^{(m)}$, $D_1^{(m)}$, $D_2^{(m)}$ are properly defined by the region of support of the cumulants to be estimated (see Appendix B).

Now the equations in (13) can be collected in matrix form as

$$\mathbf{C}_{y,y}\tilde{\mathbf{w}}_i=\mathbf{c}_{a,y}^{(i)} \qquad (14)$$

where $\mathbf{C}_{y,y}$ is a $KL \times KL$ matrix given by

$$\mathbf{C}_{y,y} = \begin{bmatrix} \mathbf{C}_{y,y}^{1,1} & \mathbf{C}_{y,y}^{1,2} & \cdots & \mathbf{C}_{y,y}^{1,K} \\ \mathbf{C}_{y,y}^{2,1} & \mathbf{C}_{y,y}^{2,2} & \cdots & \mathbf{C}_{y,y}^{2,K} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{C}_{y,y}^{K,1} & \mathbf{C}_{y,y}^{K,2} & \cdots & \mathbf{C}_{y,y}^{K,K} \end{bmatrix}$$

$$\mathbf{C}_{y,y}^{i,l} = \begin{bmatrix} c_{y,y}^{(i,l)}(0) & \cdots & c_{y,y}^{(i,l)}(L_1 - L_2) \\ c_{y,y}^{(i,l)}(1) & \cdots & c_{y,y}^{(i,l)}(L_1 - L_2 + 1) \\ \cdots & \cdots & \cdots \\ c_{y,y}^{(i,l)}(-L_1 + L_2) & \cdots & c_{y,y}^{(i,l)}(0) \end{bmatrix}$$

and $\mathbf{c}_{a,y}^{(i)}$ is a $KL \times 1$ vector given by

$$\mathbf{c}_{a,y}^{(i)} = \left[\mathbf{c}_{a,y}^{(i,1)^T} \ \ \mathbf{c}_{a,y}^{(i,2)^T} \ \ \cdots \ \ \mathbf{c}_{a,y}^{(i,K)^T}\right]^T$$

$$\mathbf{c}_{a,y}^{(i,l)} = \left[c_{a,y}^{(i,l)}(L_1),\, c_{a,y}^{(i,l)}(L_1+1),\, \cdots,\, c_{a,y}^{(i,l)}(L_2)\right]^T.$$

The matrix $\mathbf{C}_{y,y}$ is nonsingular due to **AS1**, **AS2**, **AS3**, and the fact that $\mathcal{K} \neq 0$. Some remarks regarding this statement are reported in Appendix C.

IV. Adaptive Implementation

We can consider a recursive solution of the linear system (14) if we adopt the cumulant estimation procedure reported in Appendix D. Define as $\mathbf{C}_{y,y}(n)$, $\tilde{\mathbf{w}}_i(n)$, and $\mathbf{c}_{a,y}^{(i)}(n)$ as the estimation of $\mathbf{C}_{y,y}$, $\tilde{\mathbf{w}}_i$, and $\mathbf{c}_{a,y}^{(i)}$, respectively, at time instant n. Since the estimation of fourth-order cumulants by sample statistics generally gives higher error variance than traditional correlations, the least squares solution of the linear system (14) may be ill-conditioned as a result. This motivated the choice of a particularly powerful tool for the solution of ill-conditioned problems such as total least squares (TLS) [22].

The basic idea in TLS is based on the assumption that *both* the matrix $\mathbf{C}_{y,y}(n)$ and the vector $\mathbf{c}_{a,y}^{(i)}(n)$ are subject to errors (perturbations), say $\mathbf{C}_{y,y}(n)^\epsilon$ and $\mathbf{c}_{a,y}^{(i)}(n)^\epsilon$. The computational problem to solve becomes[4]

$$\min_{\tilde{\mathbf{w}}_i(n)} \left\| [\mathbf{C}_{y,y}(n) \ \ \mathbf{c}_{a,y}^{(i)}(n)] - [\mathbf{C}_{y,y}(n)^\epsilon \ \ \mathbf{c}_{a,y}^{(i)}(n)^\epsilon] \right\|_F^2$$
$$\texttt{subject to } \mathbf{C}_{y,y}(n)^\epsilon \tilde{\mathbf{w}}_i(n) = \mathbf{c}_{a,y}^{(i)}(n)^\epsilon.$$

One way of solving problem (15) is by singular value decomposition (SVD) [7].

- Compute the SVD

$$\left[\mathbf{C}_{y,y}(n) \ \ \mathbf{c}_{a,y}^{(i)}(n)\right] = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \sum_{l=1}^{KL+1} \mathbf{u}_l \sigma_l \mathbf{v}_l^H$$

  where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{KL+1}$,
- Determine the largest integer p such that
  a) $\sigma_p > \sigma_{p+1}$;
  b) $[\mathbf{V}]_{KL+1,l} \neq 0$ for $l = p+1, p+2, \cdots, KL+1$.
- Partition $\mathbf{V}$ as $\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$ where $\mathbf{V}_{11}$ is $KL \times p$, $\mathbf{V}_{12}$ is $KL \times KL - p + 1$, $\mathbf{V}_{21}$ is $1 \times p$, and $\mathbf{V}_{22}$ is $1 \times KL - p + 1$.
- The minimum-norm TLS solution is

$$\tilde{\mathbf{w}}_i(n) = -\mathbf{V}_{12}\mathbf{V}_{22}^\dagger = -\mathbf{V}_{12}\mathbf{V}_{22}^H\left(\mathbf{V}_{22}\mathbf{V}_{22}^H\right)^{-1}. \quad (15)$$

This algorithm includes the extensions from [22].[5] The required SVD in the TLS solution makes the solution prohibitive in terms of computational complexity in real-time applications.

However, SVD can be approximated using a QR factorization followed by a Jacobi-type diagonalization procedure based on an SVD-update scheme inspired by [18]. Denote $[\mathbf{C}_{y,y}(n)]_l$ as the lth row of $\mathbf{C}_{y,y}(n)$. Similarly, $[\mathbf{c}_{a,y}^{(i)}(n)]_l$ is the lth element of $\mathbf{c}_{a,y}^{(i)}(n)$.

The *approximated* SVD (which is actually a URV decomposition) at step n is $\mathbf{U}(n)\mathbf{R}(n)\mathbf{V}^H(n)$, where $\mathbf{R}(n)$ is an upper triangular and *almost* diagonal matrix. After appending a set of new data rows to this factorization, the problem is to obtain the new decomposition $\mathbf{U}(n+1)\mathbf{R}(n+1)\mathbf{V}^H(n+1)$. This can be performed by means of the following steps.

- $\mathbf{R}(n+1) \Leftarrow \lambda \mathbf{R}(n)$.
- $\mathbf{V}(n+1) \Leftarrow \mathbf{V}(n)$.
- *for* $l = 1, 2, \cdots, KL$ *do*
  1) incorporate the new cumulants estimates

$$[\tilde{\mathbf{r}}^T \tilde{d}] = [[\mathbf{C}_{y,y}(n+1)]_l [\mathbf{c}_{a,y}^{(i)}(n+1)]_l] \mathbf{V}(n+1),$$

  2) update the QR factors:

$$\begin{bmatrix} \hat{\mathbf{R}}(n+1) \\ [\mathbf{0}^T \ \ 0] \end{bmatrix} \Leftarrow \mathbf{Q}(n+1)^H \begin{bmatrix} \mathbf{R}(n+1) \\ [\tilde{\mathbf{r}}^T \ \ \tilde{d}] \end{bmatrix}$$

  3) diagonalize by $\mathcal{N}$-step iterative plane rotations:
  for $\nu = 1, \cdots, \mathcal{N}$ *do*

$$\mathbf{R}(n+1) \Leftarrow \mathbf{P}_{1,2}\mathbf{\Theta}_{1,n,l}^H \hat{\mathbf{R}}(n+1)\mathbf{\Phi}_{1,n,l}\mathbf{P}_{1,2}$$
$$\mathbf{V}(n+1) \Leftarrow \mathbf{V}(n)\mathbf{\Phi}_{1,n,l}\mathbf{P}_{1,2}$$

  for $\kappa = 2, \cdots, KL$ *do*

$$\mathbf{R}(n+1) \Leftarrow \mathbf{P}_{\kappa,\kappa+1}\mathbf{\Theta}_{\kappa,n,l}^H \mathbf{R}(n+1)\mathbf{\Phi}_{\kappa,n,l}\mathbf{P}_{\kappa,\kappa+1}$$
$$\mathbf{V}(n+1) \Leftarrow \mathbf{V}(n+1)\mathbf{\Phi}_{\kappa,n,l}\mathbf{P}_{\kappa,\kappa+1}$$

  end
  end
- *end.*

[4] We use the notation $\|\mathbf{v}\| = \sqrt{\sum_{i=1}^M |v_i|^2}$ for the 2-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 $m_{i,j}$.

[5] The conditions in Step 2 ensure that the algorithm computes the unique minimum-norm solution and that such a solution exists. Observe also that

$$\|\tilde{\mathbf{w}}_i(n)\| = \sqrt{\|\mathbf{V}_{22}\|^{-2} - 1}$$

and

$$\|[\mathbf{C}_{y,y}(n) \ \ \mathbf{c}_{a,y}^{(i)}(n)] - [\mathbf{C}_{y,y}(n)^\epsilon \ \ \mathbf{c}_{a,y}^{(i)}(n)^\epsilon]\|_F = \sqrt{\sum_{l=p+1}^{KL+1} \sigma_l^2}.$$

Whenever $p = KL$, the full TLS solution is obtained. If $p < KL$, the so-called truncated TLS (T-TLS) solution is obtained.

Initialization is $\mathbf{V}(0) \Leftarrow \mathbf{I}_{KL}$, and $\mathbf{R}(0) \Leftarrow \mathbf{0}_{KL}$, where $\mathbf{I}_{KL}$ is a $KL \times KL$ identity matrix and $\mathbf{0}_{KL}$ is a $KL \times KL$ matrix whose entries are all zero. At this point partitioning $\mathbf{V}(n+1)$ as $\mathbf{V}(n+1) = \begin{bmatrix} \mathbf{V}_{11}(n+1) & \mathbf{V}_{12}(n+1) \\ \mathbf{V}_{21}(n+1) & \mathbf{V}_{22}(n+1) \end{bmatrix}$, it is possible to obtain the minimum-norm TLS solution for $\tilde{\mathbf{w}}_i(n+1)$ using (15). Observe that the $\mathbf{U}$-factor of the factorization is not propagated. In Step 2), the QR factorization update [7] is performed by means of a sweep of the new input data. The plane rotations $\mathbf{\Phi}_{\kappa,n,l}$ and $\mathbf{\Theta}_{\kappa,n,l}$ of Step 3) are unitary transformations chosen as to annihilate the $(\kappa, \kappa+1)$ element of $\mathbf{R}(n+1)$ [18], [20], while leaving it upper triangular.[6] The matrices $\mathbf{P}_{\kappa,\kappa+1}$ are permutation matrices in the $(\kappa, \kappa+1)$ plane

$$\mathbf{P}_{\kappa,\kappa+1} = \begin{bmatrix} \mathbf{I}_{\kappa-1} & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & \mathbf{I}_{KL-\kappa} \end{bmatrix}.$$

Stewart in [20] was the first one to observe that two-sided orthogonal transformations of this type iteratively applied to a matrix will reduce the R-factor to almost diagonal, and the diagonal elements will approximate the singular values [18]. The number of iterations one has to perform to obtain the desired degree of accuracy is indicated as $\mathcal{N}$. In applications where computational efficiency is of concern, $\mathcal{N}$ must be as low as one. In extremely fast-fading environments where tracking capability of the algorithm is top priority, the value of $\mathcal{N}$ should be in the range of three to five. In any case, the higher the value of $\mathcal{N}$, the closer the SVD is approximated at every step. It is important to observe that several successful refinements of this elegant algorithm were proposed by the authors of [18]. To detail these refinements is out of the scope of this work. One of the problems revealed in [18] is the fact that the matrix $\mathbf{V}(n)$ as computed by accumulation of Jacobi rotation matrices gradually loses orthogonality. Our application of the method, however, never justified (in terms of performance improvement) the increased computational complexity subsequent to the orthogonalization procedure proposed in [18]. One of the reasons is the bursty nature of the TDMA system: the receiver never processes more than 500–1000 samples (depending on the sampling rate), and it appears that the deviation from orthogonality becomes consistent for larger sample sizes. However, the algorithm is very effective in tracking slow variations of the data matrices, and more important, is much more stable than a full-RLS solution. The value of the scheme is in the fact that the complexity per time step is an $\mathcal{O}((KL+1)^2)$ if $\mathcal{N} = 1$. Moreover, including orthogonalization of the V-matrix produces an algorithm with high accuracy (similar to the cyclic Jacobi algorithm for SVD [9]).

[6] The plane rotations $\mathbf{\Phi}_{\kappa,n,l}$ and $\mathbf{\Theta}_{\kappa,n,l}$ of Step 3) have the following structure

$$\mathbf{\Theta}_{\kappa,n,l} = \begin{bmatrix} \mathbf{I}_{\kappa-1} & & & \\ & \cos\theta_{\kappa,n,l} & r_{\kappa,n,l}\sin\theta_{\kappa,n,l} & \\ & -r^*_{\kappa,n,l}\sin\theta_{\kappa,n,l} & \cos\theta_{\kappa,n,l} & \\ & & & \mathbf{I}_{KL-\kappa} \end{bmatrix}$$

$$\mathbf{\Phi}_{\kappa,n,l} = \begin{bmatrix} \mathbf{I}_{\kappa-1} & & & \\ & \cos\psi_{\kappa,n,l} & s_{\kappa,n,l}\sin\psi_{\kappa,n,l} & \\ & -s^*_{\kappa,n,l}\sin\psi_{\kappa,n,l} & \cos\psi_{\kappa,n,l} & \\ & & & \mathbf{I}_{KL-\kappa} \end{bmatrix}$$

where $\mathbf{I}_n$ is a $n \times n$ identity matrix and $r_{\kappa,n,l}$, $s_{\kappa,n,l}$, $\theta_{\kappa,n,l}$, $\psi_{\kappa,n,l}$ are obtained from $\mathbf{R}(n+1)$ (see [7], [9], [18], and [20] for more details).

V. Results of Experiments

A TDMA system for cellular communications has been simulated according to [21]. In addition, we present the results of some lab experiments collected using the Watkins–Johnson wideband dual-mode (AMPS and IS-136) base-station $Base_2$. A block diagram of the receiver section of the base-station is shown in Fig. 1. The tuner module performs a standard single conversion scheme. The A/D is a high-speed bandpass sampler, while the conversion at baseband is operated by digital downconverters (wideband processing). Of particular importance is the fact that we always compare the proposed approach (HOS-TLS) with a more traditional QR-RLS (recursive least squares based on QR decomposition [9]) approach, explicitly a second-order statistics method.

A. Remarks on the Implementation

The simulations in the following section are performed using a wordlength size of 16 bits, using fixed-point arithmetic. Digital signal processing processors are commercially available with these characteristics. The computational complexity in terms of multiplications, square roots, and reciprocals per iteration was calculated and compared to the complexity of the adaptive QR-RLS. The approximate number of instructions per second using filter lengths equal to $L = 7$ and $K = 2$, including open-loop synchronization, frequency offset compensation, and convolutional decoding of the bitstream, is expressed in million of instructions per second (MIPS) per time-slot (full-rate channel)

$$\text{QR-RLS} \Rightarrow 32 \text{ MIPS}$$
$$\text{HOS-TLS } (\mathcal{N} = 1) \Rightarrow 48 \text{ MIPS}.$$

Square-roots are performed using series approximation, divisions use the analog devices engine. It is evident that the HOS algorithm proposed here has considerably higher computational complexity than the traditional QR-RLS scheme, but it is important to emphasize that the significant improvement in performance may justify the choice.

B. Computer Simulations Results

In the simulations we assumed a sensor spacing equal $\tilde{\lambda}/2$. The antenna has five elements and there are three transmitting mobiles. One of the transmitters is at array broadside and is the signal of interest. We assume a two-ray model, and each path impulse response is modeled as a two-path Rayleigh fading channel $(P_1^{(1)} = P_2^{(1)} = 2)$ [21]. The arrival angles of the two paths are spread around $\theta^1 = 0°$ with a cluster width of $2°$ (this model refers to the geometric interpretation also used in [1]). The interfering signals are generated with the same parameters but with DOA's clustered around $\theta^2 = 30°$ and $\theta^3 = -70°$. In Fig. 2, the equivalent baseband discrete-time model is shown relative to the ith mobile transmitter. Delay spread propagation parameters are summarized in Table I 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

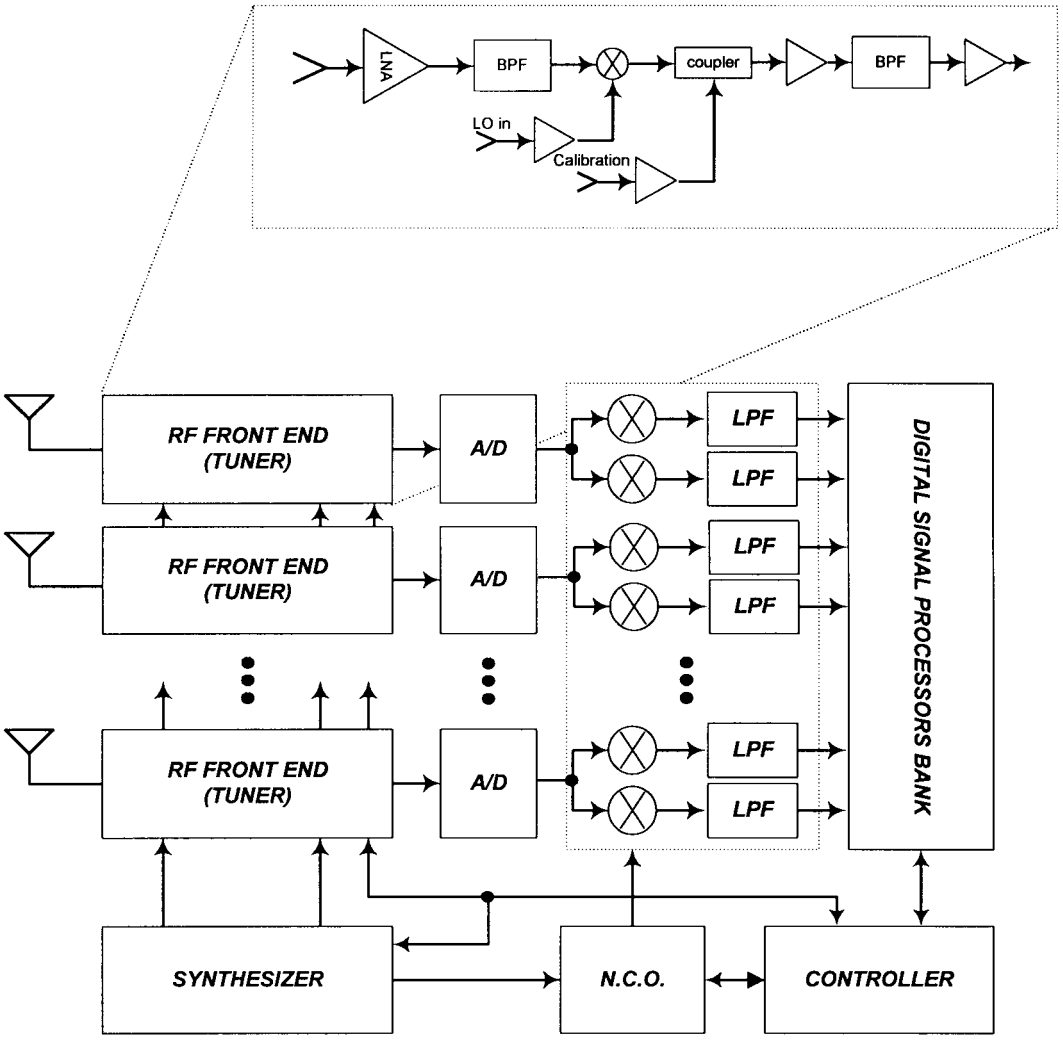


Fig. 1. Block diagram of the receiver.

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. In particular, each process has a frequency response equal to the square-root of the Doppler power density spectrum.[7] Table II summarizes the Doppler frequency situation as related to the test cases results shown in the figures. The MSE is defined as the average of the squared error obtained over $\tilde{M} = 100$ Monte Carlo runs and is given by $\text{MSE}_1 = (1/\tilde{M}) \sum_{\nu=1}^{\tilde{M}} |\epsilon_i^{(\nu)}(n)|^2$. The error obtained at the mth run is $\epsilon_i^{(\nu)}(n) = z_i^{(\nu)}(n) - a_i(n - n_i)$ where n_i is the delay introduced by the filters and $z_i^{(\nu)}(n)$ is the output of the combined filters obtained at the νth run relative to the ith transmitter. Observe that the proposed algorithm uses the known synchronization sequence while in *training mode* and past decisions while in *decision directed* mode. The slots are 162 symbols long, and 14 symbols at the beginning of each slot are known at the receiver.[8] The SNR for each discrete-time channel impulse response is defined as in [2]. In Fig. 3, the traces of the filtering stage relative to the first user are shown using particular test slots for $\text{SNR}_{bit} = 21$ dB. Observe that the filter after the first 14 symbols is running in decision-directed mode. The QR-RLS algorithm on the left is compared with the proposed approach. The plots represent the traces of the four largest magnitudes among the taps of the three five-input one-input filters of length $L = 7$. Fig. 4 shows an experiment where the synchronization sequences are assumed to be infinitely long (that is, the receiver has perfect knowledge of the transmitted information). The plots show the trajectory of the magnitude of the weights as compared to the optimum Wiener filter (constrained to have finite length) computed assuming perfect instantaneous knowledge of $\mathcal{H}(z)$. The forgetting factors are $\lambda = 0.97, 0.9$, and 0.85, respectively, in the plots and $\text{SNR}_{bit} = 30$ dB. Fig. 5 shows the MSE averaged over 100 independent computer runs versus time-step in symbols when using $L = 7$ for

[7] 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.

[8] This corresponds to a frame format similar to the data traffic channel (DTC) of [21].

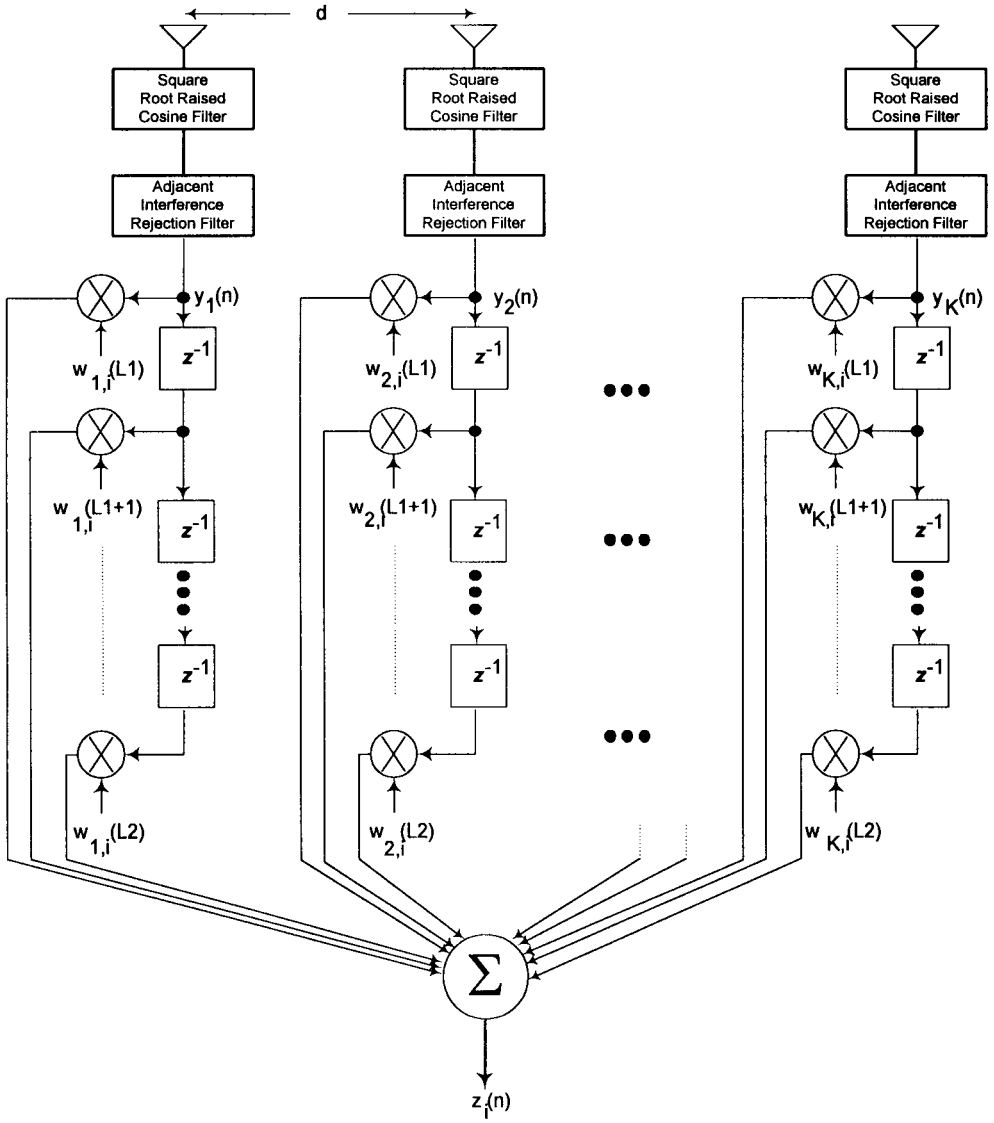


Fig. 2. Discrete-time model of the filtering section (K sensors) relative to the ith transmitter. Observe that an adjacent interference-rejection filter is concatenated with the (square-root) raised-cosine filter which by itself does not meet the IS-136 specification in terms of out-of-band rejection.

TABLE I
CHANNEL PROPAGATION ENVIRONMENTS FOR PERFORMANCE EVALUATION RESULTS: DOA'S AND DELAY SPREADS OF THE MOBILES. IN ALL THE 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	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. 3	5	3	30°, −70°	41.2 μsec	41.2 μsec	10.3 μsec	10.3 μsec	5.15 μsec	5.15 μsec
Fig. 4	3	2	30°	10.3 μsec	10.3 μsec	0.0 μsec	0.0 μsec	− − . − −	− − . − −
Fig. 5	5	3	30°, −70°	41.2 μsec	41.2 μsec	10.3 μsec	10.3 μsec	10.3 μsec	10.3 μsec
Fig. 6	5	3	30°, −70°	30.9 μsec	30.9 μsec	10.3 μsec	10.3 μsec	20.6 μsec	20.6 μsec
Fig. 8	2	1	− − . − −	41.2 μsec	− − . − −	− − . − −	− − . − −	− − . − −	− − . − −

$\text{SNR}_{bit} = 21$ dB. Bit error rate analysis results are shown in Fig. 6 (at different speeds) with delays as specified in Table I. The length of the filters is $L = 7$, $L_1 = -2$, and $L_2 = 4$. The forgetting factor is $\lambda = 0.97$. The results are compared with the QR-RLS. The SNR is the same on each discrete-time channel; that is, all co-channel interferers are received with the same average power. A sample size of 10^{k+4} was used to estimate an error probability of 10^{-k}.

C. Hardware Implementation Results

A simpler and indeed more realistic scenario ($K = 2$, $U = 1$) was studied using data collected from the DSP receiver section of $Base_2$, the dual-mode wideband base station implemented at Watkins–Johnson. The hardware test setup is depicted in Fig. 7. A hardware multipath fading simulator is connected to the two antenna ports of the base station. The

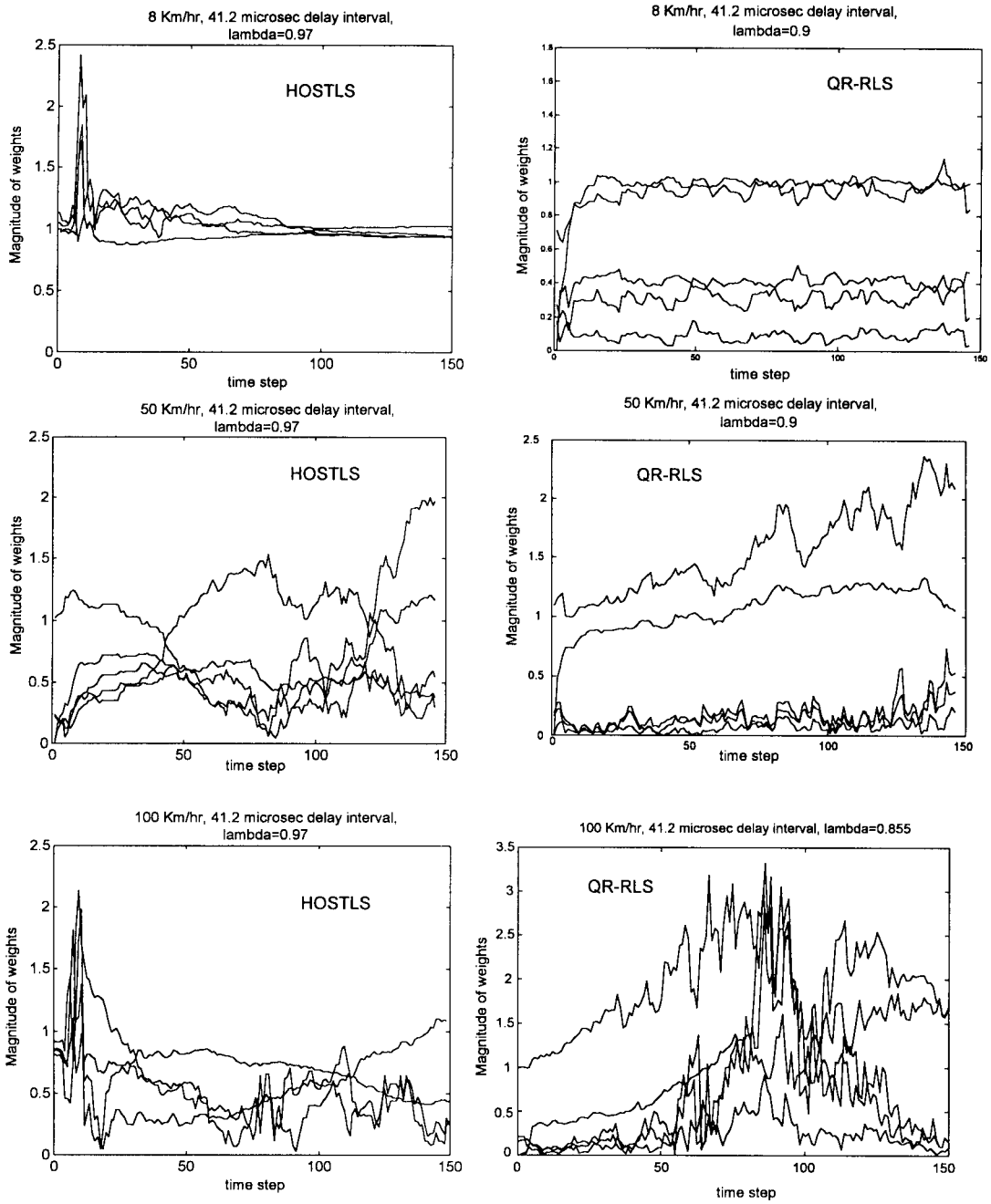


Fig. 3. Trajectories of the magnitude of the filter weights corresponding to the first mobile for different speeds. The length of the filter is $L = 7$. For propagation parameters, see Tables I and II.

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

	V_1	V_2	V_3
Fig. 3	$8/50/100\ Km/hr$	$50\ Km/hr$	$50\ Km/hr$
Fig. 4	$100\ Km/hr$	$100\ Km/hr$	$---$
Fig. 5	$50/100\ Km/hr$	$100\ Km/hr$	$50\ Km/hr$
Fig. 6	$100\ Km/hr$	$100\ Km/hr$	$50\ Km/hr$
Fig. 8	$8/50/100\ Km/hr$	$---$	$---$

IS-136 signal generator simulates transmission of DTC frames coming from three different mobiles. Additive Gaussian noise is injected on both the diversity channels. Observe that the DSP modem receives a sampling rate of 80 KHz. The wordlength used is 16 and the algorithm has been implemented using fixed-point arithmetic.[9] A polyphase raised-cosine filter concatenated with an adjacent interference rejection filter transforms the rate to $4/T = 97.2$ kHz, which is the rate at which the open-loop synchronizer works. Then the two-channel filter works at $2/T$-rate. The results of extensive BER measurements are summarized in Fig. 8. The QR-RLS algorithm opposes unsatisfactory performance in many situations. It is important to mention, however, that substantial improvement can be

[9] A detailed analysis of the dynamic range required not to degrade the performance of the algorithm as opposed to the floating point representation was carried out although the description of such analysis is beyond the scope of this paper.

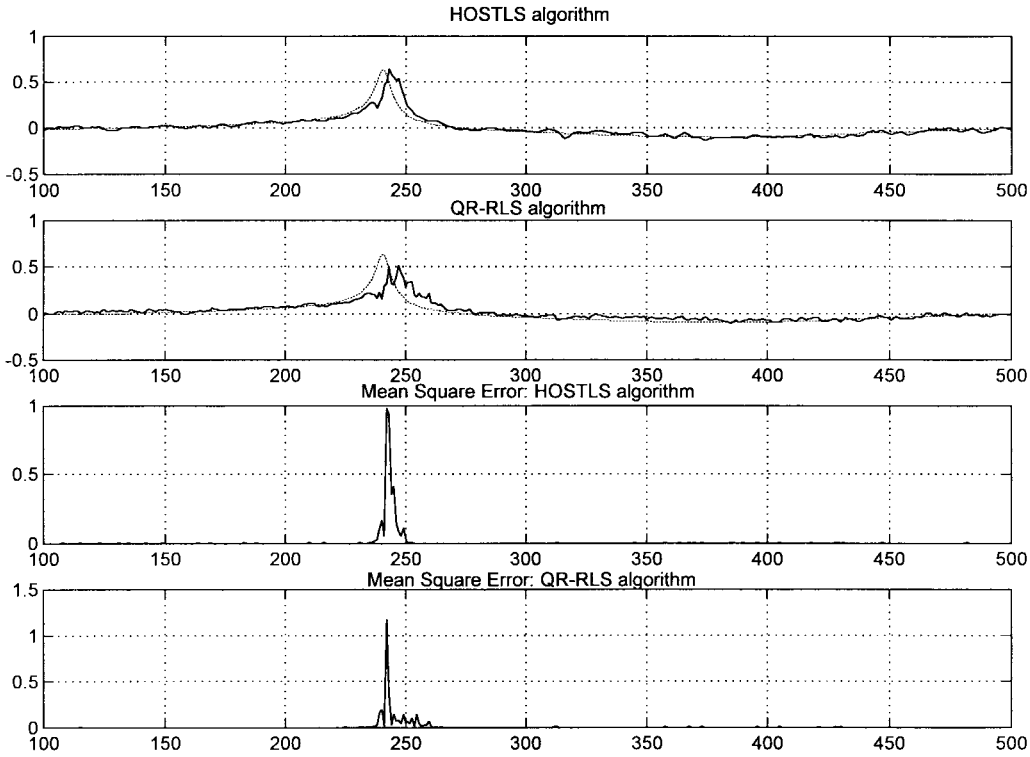


Fig. 4. An experiment that shows the tracking capability of the algorithm. The dashed trace is the optimum Wiener filter computed assuming perfect knowledge of the discrete-time of the channel impulse response. The length of the filter is $L = 3$. For propagation parameters, see Tables I and II.

achieved using QR-based decision-feedback schemes, and similar extensions of the cumulant-based algorithm are possible.

VI. Conclusions

We have studied a new solution to the array processing problem in a cellular TDMA base station employing antenna arrays. The method is robust and able to track fast channel variations caused by moving transmitters. Although the algorithm exploits higher order statistics which usually are employed in blind algorithms with relatively low convergence speed, the knowledge of the training/synchronization sequences allowed a formulation of the problem as a known-input identification/deconvolution problem and gave the possibility of designing a method with fast convergence. The MSE of traditional second-order statistics algorithms has been shown to be proportional to a cost function involving cumulants of the restoration error. From this basic observation we derived a set of linear equations relating the separating filters to input–output fourth-order cumulants. The well-known high variance exhibited by short data record estimates of cumulants is mitigated by the use in the adaptive implementation of a recursive total least squares approach. Results for the IS-136 (the U.S. standard for digital cellular communications) have been shown as they compare with more traditional algorithms based on second-order statistics.

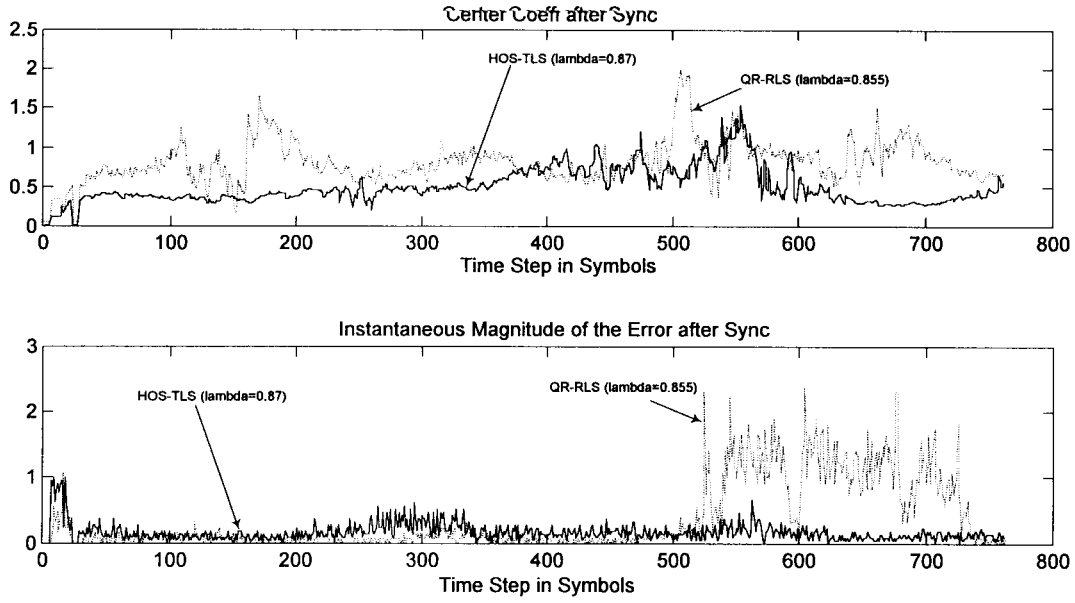
Appendix A
Proof of Theorem 2

Suppose $w^+_{i,l}(n)$ are the filter weights organized into $\tilde{\mathbf{w}}^+_i$ satisfying (11) and (12) and minimizing $J^{(i)}(\tilde{\mathbf{w}}^+_i)$, while $w_{i,l}(n)$ organized into $\tilde{\mathbf{w}}_i$ are the weights of another arbitrary filter. Define $\epsilon^+_i(n) = a_i(n) - \sum_l w^+_{i,l}(n) * y_l(n)$ and observe that it is possible to express $\epsilon_i(n)$ as

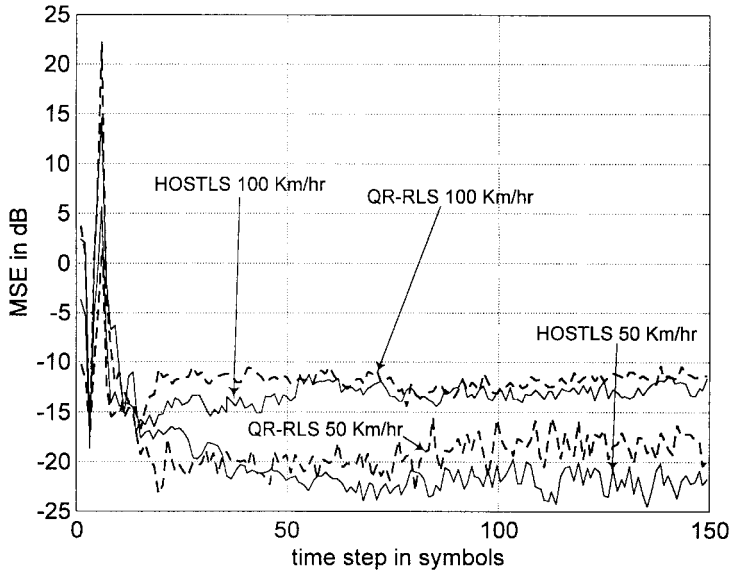
$$\begin{aligned}\epsilon_i(n) &= \left[a_i(n) - \tilde{\mathbf{w}}_i^{+^T}\tilde{\mathbf{y}}(n)\right] + (\tilde{\mathbf{w}}^+_i - \tilde{\mathbf{w}}_i)^T\tilde{\mathbf{y}}(n)\\ &= \epsilon^+_i(n) + (\tilde{\mathbf{w}}^+_i - \tilde{\mathbf{w}}_i)^T\tilde{\mathbf{y}}(n) \qquad (16)\end{aligned}$$

where $\tilde{\mathbf{y}}(n) = [\tilde{\mathbf{y}}_1^T(n), \tilde{\mathbf{y}}_2^T(n), \cdots, \tilde{\mathbf{y}}_K^T(n)]^T$, $\tilde{\mathbf{y}}_l(n) = [y_l(n - L_1), \cdots, y_l(n - L_2)]^T$. Substituting (16) into the expression for $J^{(i)}(\tilde{\mathbf{w}}_i)$ and using (11) and (12), it is possible to write

$$\begin{aligned}&J^{(i)}(\tilde{\mathbf{w}}_i)\\ &= \left|\sum_{l_1}\sum_k \operatorname{cum}\left[\epsilon^+_i(n), \epsilon^{+^*}_i(n), y_{l_1}(n-k), y^*_{l_1}(n-k)\right]\right.\\ &\quad + \sum_{l_1}\sum_{l_2}\sum_{m_1}\sum_{m_2}\left[w^+_{i,l_1}(m_1) - w_{i,l_1}(m_1)\right]\\ &\quad \times \left[w^+_{i,l_2}(m_2) - w_{i,l_2}(m_2)\right]^*\\ &\quad \left.\times \mathcal{K}\sum_{l_3}\sum_n g_{l_3,l_1}(n-m_1)g^*_{l_3,l_2}(n-m_2)\right|^2\\ &= \left|\frac{\mathcal{K}}{\sigma_x^2}\langle E\{|\epsilon^+_i(n)|^2\}\rangle_{\mathrm{SNR}=\infty}\right.\\ &\quad \left. + \mathcal{K}\sum_l\sum_n |(\tilde{\mathbf{w}}^+_i - \tilde{\mathbf{w}}_i)\mathbf{g}_l(n)|^2\right|^2\\ &= \left|\frac{\mathcal{K}}{\sigma_x^2}\right|^2\left|\langle E\{|\epsilon^+_i(n)|^2\}\rangle_{\mathrm{SNR}=\infty}\right.\\ &\quad \left. + \sigma_x^2\sum_l\sum_n |(\tilde{\mathbf{w}}^+_i - \tilde{\mathbf{w}}_i)\mathbf{g}_l(n)|^2\right|^2 \qquad (17)\end{aligned}$$



(a)



(b)

Fig. 5. Stability experiments (a) for the two algorithms: QR-RLS and HOS-TLS at 100 km/h. The slot is 750 symbols (for testing purposes only). (b) Averaged mean squared error (100 Monte Carlo runs). For propagation parameters, see Tables I and II.

where

$$\mathbf{g}_i(n) = \left[\mathbf{g}_{i,1}^T(n), \mathbf{g}_{i,2}^T(n), \cdots, \mathbf{g}_{i,K}^T(n)\right]^T$$
$$\mathbf{g}_{i,l}(n) = [g_{i,l}(n-L_1), \cdots, g_{i,l}(n-L_2)]^T.$$

Of course, (17) is minimized when $\tilde{\mathbf{w}}_i^+ = \tilde{\mathbf{w}}_i$. Q.E.D.

APPENDIX B
ON THE CHOICE OF $D_1^{(m)}$, $B_1^{(m)}$, $D_2^{(m)}$, $B_2^{(m)}$

Assume $g_{i,l}(n)$ are FIR of length $r = r_2 - r_1 + 1$ and observe that

$$\begin{aligned}&\operatorname{cum}\left[y_i(n), y_{i_1}^*(n-m), y_l(n-k), y_l^*(n-k)\right]\\ &= \operatorname{cum}\left[y_i(n+k), y_{i_1}^*(n+k-m), y_l(n), y_l^*(n)\right]\\ &= \gamma_{4x}\sum_{l_1}\sum_n g_{i,l_1}(n+k)g_{i_1,l_1}^*(n+k-m)|g_{l,l_1}(n)|^2\end{aligned}$$

is different from zero only for

$$\max\{-r+1, -r+1+m\} \le k \le \min\{r-1, r-1+m\}$$

so that $B_1^{(m)}$ and $B_2^{(m)}$ can be chosen as

$$\begin{aligned}B_1^{(m)} &\le \max\{-r+1, -r+1+m\}\\ B_2^{(m)} &\ge \min\{r-1, r-1+m\}.\end{aligned}$$

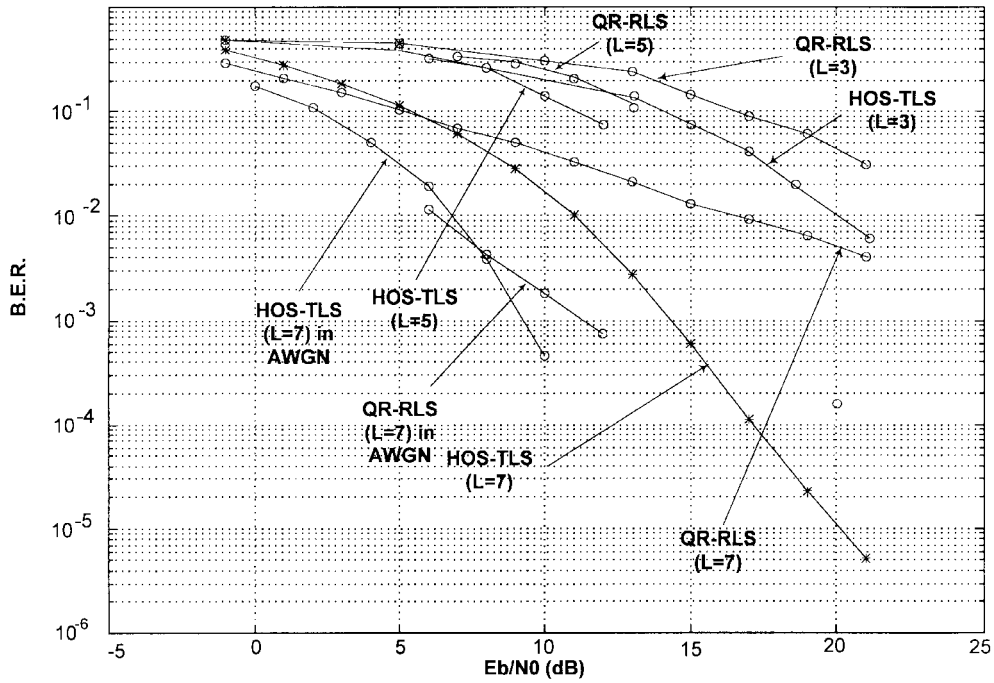


Fig. 6. Bit error rate of the first mobile transmitter in a three-mobile environment. For propagation parameters, see Tables I and II. Also, results for the no-fading case are shown.

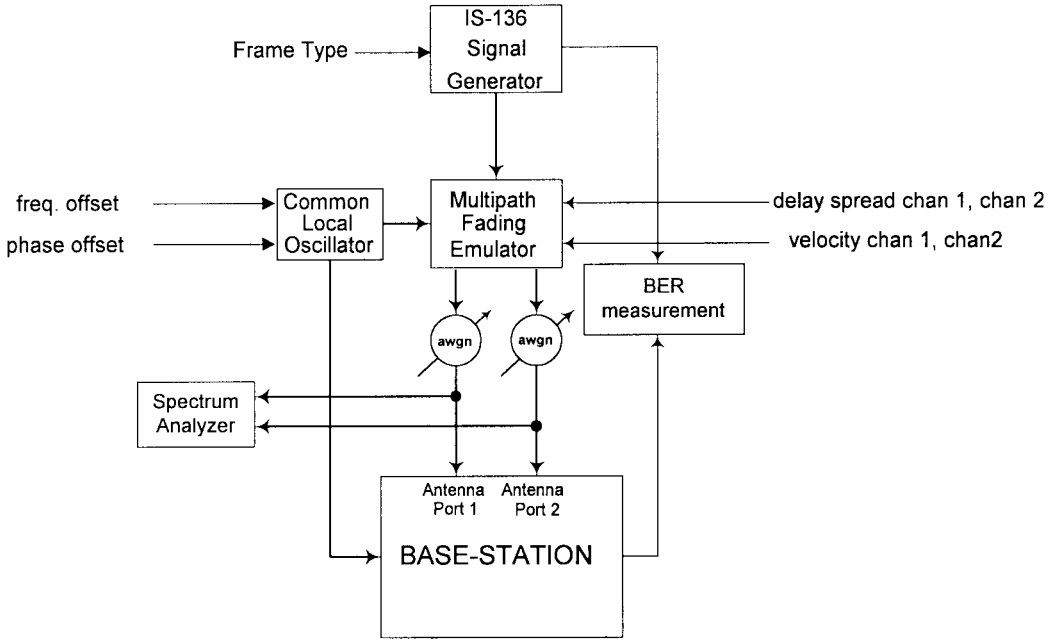


Fig. 7. Hardware test setup for lab experiments.

Remember now that $b_i(n)$ (see **AS3**) is of length $M = M_2 - M_1 + 1$ and observe that

$$\begin{aligned}&\operatorname{cum}\left[a_i(n),\, y_{i_1}^*(n-m),\, y_l(n-k),\, y_l^*(n-k)\right]\\&\quad=\operatorname{cum}\left[a_i(n+k),\, y_{i_1}^*(n+k-m),\, y_l(n),\, y_l^*(n)\right]\\&\quad=\gamma_{4x}\sum_{l_1}\sum_{n} b_i(n+k)g_{i_1,l_1}^*(n+k-m)|g_{l,l_1}(n)|^2\end{aligned}$$

is different from zero only for $M_1 \le n+k \le M_2$, $r_1 \le n+k-m \le r_2$, and $r_1 \le n \le r_2$. This means also that $M_1 - n \le k \le M_2 - n$, $r_1 - n + m \le k \le r_2 - n + m$, $r_1 \le n \le r_2$,

$$\max\{M_1-n,\, r_1-n+m\} \le k \le \min\{M_2-n,\, M_2-n+m\},$$

and $r_1 \le n \le r_2$, that is,

$$\max\{M_1-r_2,\, r_1-r_2+m\} \le k \le \min\{M_2-r_1,\, r_2-r_1+m\}.$$

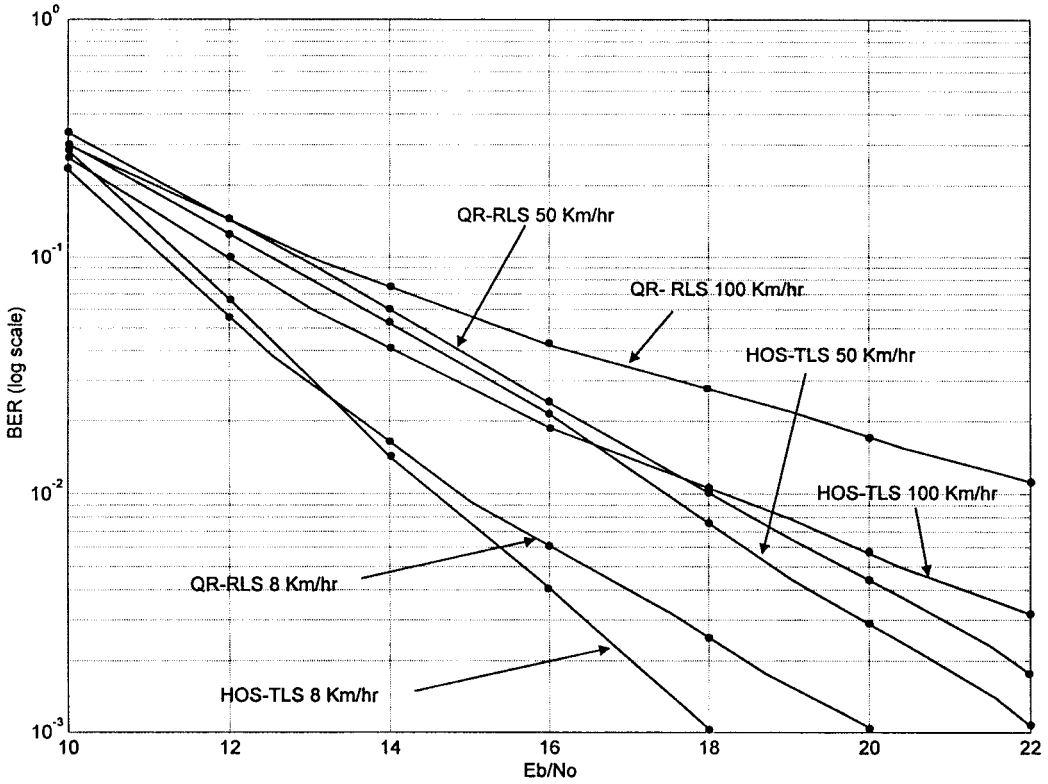


Fig. 8. Bit error rate for the hardware experiments. For propagation parameters, see Tables I and II. Observe that fixed-point effects (16 bits) are deteriorating the performance with respect to the results shown in Fig. 6.

Hence $D_1^{(m)}$ and $D_2^{(m)}$ can be chosen as

$$D_1^{(m)} \le \max\{M_1 - r_2, -M+1+m\}$$
$$D_2^{(m)} \ge \min\{M_2 - r_1, r-1+m\}.$$

Appendix C
On the Rank of the Matrix $\mathbf{C}_{y,y}$

Define the $K(M+J+L-2)\times KL$ matrix

$$\tilde{\mathbf{G}} = \begin{bmatrix} \mathbf{G}_{1,1} & \mathbf{G}_{2,1} & \cdots & \mathbf{G}_{K,1} \\ \mathbf{G}_{1,2} & \mathbf{G}_{2,2} & \cdots & \mathbf{G}_{K,2} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{G}_{1,U} & \mathbf{G}_{2,U} & \cdots & \mathbf{G}_{K,U} \end{bmatrix}$$

$$[\mathbf{G}_{i,k}]_{m,n} = \begin{cases} g_{i,k}(m-n), & M_1+J_1 \le m-n \\ & \le M_2+J_2 \\ 0, & \text{elsewhere} \end{cases}$$
$$M_1+J_1+L_1 \le m \le M_2+J_2+L_2,$$
$$L_1 \le n \le L_2$$

the $K(J+L-1)\times KL$ matrix

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

$$[\mathbf{H}_{i,k}]_{m,n} = \begin{cases} h_{i,k}(m-n), & J_1 \le m-n \le J_2 \\ 0, & \text{elsewhere} \end{cases}$$
$$J_1+L_1 \le m \le J_2+L_2,\ L_1 \le n \le L_2$$

and the $K(M+J+L-2)\times K(J+L-1)$ matrix

$$\tilde{\mathbf{B}} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_U \end{bmatrix}$$

$$[\mathbf{B}_i]_{m,n} = \begin{cases} b_i(m-n), & M_1 \le m-n \le M_2 \\ 0, & \text{elsewhere} \end{cases}$$
$$M_1+J_1+L_1 \le m \le M_2+J_2+L_2,$$
$$J_1+L_1 \le n \le J_2+L_2.$$

Now observe that

$$\begin{aligned} c_{y,y}^{(i_1,i_2)}(m_2-m_1) &= \sum_{l=1}^{K}\sum_{k=-\infty}^{\infty} \operatorname{cum}\big[y_{i_1}(n-m_1), y_{i_2}^*(n-m_2), \\ &\qquad y_l(n-k), y_l^*(n-k)\big] \\ &= \gamma_{4x}\sum_{l=1}^{K}\sum_{k=-\infty}^{\infty}\sum_{l_1=1}^{U}\sum_{n} g_{i_1,l_1}(n-m_1) \\ &\quad \times g_{i_2,l_1}^*(n-m_2)|g_{l,l_1}(n-k)|^2 \\ &= \gamma_{4x}\sum_{l=1}^{K}\sum_{k=-\infty}^{\infty}|g_{l,r}(k)|^2 \\ &\quad \times \sum_{l_1=1}^{U}\sum_{n} g_{i_1,l_1}(n-m_1)g_{i_2,l_1}^*(n-m_2) \\ &= \mathcal{K}\sum_{l_1=1}^{U}[\mathbf{G}_{i_2,l_1}^H\mathbf{G}_{i_1,l_1}]_{m_2,m_1}. \end{aligned} \tag{18}$$

The last equality states that

$$\mathbf{C}_{y,y} = \mathcal{K}\tilde{\mathbf{G}}^H\tilde{\mathbf{G}}. \tag{19}$$

The three following facts:

- $\tilde{\mathbf{H}}$ is full rank [10], [12], [13], [15] due to **AS2**;
- $\tilde{\mathbf{B}}$ is full rank due to **AS3**;
- $\tilde{\mathbf{G}} = \tilde{\mathbf{B}}\tilde{\mathbf{H}}$;

imply that $\tilde{\mathbf{G}}$ is a full (column) rank matrix and that $\tilde{\mathbf{G}}^H\tilde{\mathbf{G}}$ has rank equal to KL. Since $\mathcal{K} \neq 0$, the full rank property of $\tilde{\mathbf{G}}^H\tilde{\mathbf{G}}$ extends to $\mathbf{C}_{y,y}$ due to **AS1** and (19).

Appendix D
Cumulants Estimation

Adaptive estimation of cumulants can be implemented by means of the method also used in [8]. We define

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

and the estimates of the respective moments based on sample statistics as $\hat{m}^{a,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n)}$, $\hat{m}^{y,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n)}$, $\hat{m}^{a,y}_{i,l}(k)^{(n)}$, and $\hat{m}^{y,y}_{i,l}(k)^{(n)}$ using n samples. Assuming at iteration 0 we have available $y_i(0), \cdots, y_i(I_{lag})$ for $i = 1, 2, \cdots, K$, at iteration n we can update $\hat{m}^{a,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n)}$ from $\hat{m}^{a,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n-1)}$ as follows:

$$\begin{aligned}
&\hat{m}^{a,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n)}\\
&\quad = (1-\alpha(n))\hat{m}^{a,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n-1)} + \alpha(n)a_{i_1}\Big(\Gamma^{(n,m,k)}\Big)\\
&\qquad \times y^*_{i_2}\Big(\Gamma^{(n,m,k)} - m\Big) \times y_{i_3}\Big(\Gamma^{(n,m,k)} - k\Big)\\
&\qquad \times y^*_{i_4}\Big(\Gamma^{(n,m,k)} - k\Big)
\end{aligned}$$

and $\hat{m}^{y,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n)}$, from $\hat{m}^{y,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n-1)}$ as follows:

$$\begin{aligned}
&\hat{m}^{y,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n)}\\
&\quad = (1-\alpha(n))\hat{m}^{y,y}_{i_1,i_2,i_3,i_4}(m,k)^{(n-1)} + \alpha(n)y_{i_1}\Big(\Gamma^{(n,m,k)}\Big)\\
&\qquad \times y^*_{i_2}\Big(\Gamma^{(n,m,k)} - m\Big) \times y_{i_3}\Big(\Gamma^{(n,m,k)} - k\Big)\\
&\qquad \times y^*_{i_4}\Big(\Gamma^{(n,m,k)} - k\Big)
\end{aligned}$$

where $\alpha(n) = 1/n + I_{lag}$ and $\Gamma^{(n,m,k)} = \min(n + I_{lag}, n + I_{lag} + k, n + I_{lag} + m)$.

Similarly, for the second-order moments, we can write

$$\begin{aligned}
\hat{m}^{a,y}_{i,l}(k)^{(n)} &= (1-\alpha(n))\hat{m}^{a,y}_{i,l}(k)^{(n-1)}\\
&\quad + \alpha(n)a_i\Big(\Gamma_2^{(n,k)}\Big)y^*_l\Big(\Gamma_2^{(n,k)} - k\Big)
\end{aligned} \tag{20}$$

where $\alpha(n) = 1/n + I_{lag}$ and $\Gamma_2^{(n,k)} = \min(n + I_{lag},\, n + I_{lag} + k)$.

Evidently we can obtain the cumulants estimation at point n as

$$\begin{aligned}
\hat{c}^{i_1,i_2}_{a,y}(m)^{(n)} &= \sum_{l=1}^{K}\sum_{k=D_1^{(m)}}^{D_2^{(m)}} \Big(\hat{m}^{a,y}_{i_1,i_2,l,l}(m,k)^{(n)}\\
&\qquad - \hat{m}^{a,y}_{i_1,i_2}(m)^{(n)}\hat{m}^{y,y}_{l,l}(0)^{(n)}\\
&\qquad - \hat{m}^{a,y}_{i_1,l}(k)^{(n)}\hat{m}^{y,y}_{i_2,l}(m-k)^{(n)}\Big)\\
\hat{c}^{i_1,i_2}_{y,y}(m)^{(n)} &= \sum_{l=1}^{K}\sum_{k=B_1^{(m)}}^{B_2^{(m)}} \Big(\hat{m}^{y,y}_{i_1,i_2,l,l}(m,k)^{(n)}\\
&\qquad - \hat{m}^{y,y}_{i_1,i_2}(m)^{(n)}\hat{m}^{y,y}_{l,l}(0)^{(n)}\\
&\qquad - \hat{m}^{y,y}_{i_1,l}(k)^{(n)}\hat{m}^{y,y}_{i_2,l}(m-k)^{(n)}\Big)
\end{aligned}$$

and an obvious organization of these quantities into the matrices of the system (14) gives the estimates $\mathbf{C}_{y,y}(n)$ and $\mathbf{c}_{a,y}(n)$.

Acknowledgment

The author wishes to express his gratitude to Watkins–Johnson management for interest and support throughout the course of the investigation described in this paper and to J. Taylor, Staff Scientist, for extensive and enlightening discussions on digital radio architectures.

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. Syst. Technol.*, vol. 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] A. N. Delopoulos and G. B. Giannakis, "Strongly consistent identification algorithms and noise insensitive MSE criteria," *IEEE Trans. Signal Processing*, vol. 30, no. 8, pp. 1955–1970, Aug. 1992.
[6] C. Feng and C. Chi, "Design of Wiener filters using a cumulant-based MSE criterion," *Signal Processing*, 1996, vol. 54, pp. 23–48.
[7] G. H. Golub and C. F. Van Loan, *Matrix Computation.* North Oxford Academic Publishing Co., Johnson Hopkins Press, 1988.
[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] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, NJ: Prentice Hall, 1986.
[10] Y. Inouye and K. Hirano, "Cumulant-based blind identification of linear multi-input multi-output systems driven by colored inputs," *IEEE Trans. Signal Processing*, vol. 45, pp. 1543–1552, June 1997.
[11] T. Kailath, *Linear Systems.* Englewood Cliffs, NJ: Prentice Hall, 1980.
[12] Y. Li and K. J. Ray Liu, "On blind equalization of MIMO channels," in *ICC'96, Proc.*, pp. 1000–1004.
[13] ——, "Blind MIMO FIR channel identification based on second-order statistics with multiple signal recovery," University of Maryland, College Park, ISR TR 96-7 Tech. Rep.
[14] M. Martone, "On-board regeneration of uplink signals using a blind adaptive multichannel estimator," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 49–62, Jan. 1998.
[15] ——, "An adaptive algorithm for adaptive antenna array low-rank processing in cellular TDMA base-stations," *IEEE Trans. Commun.*, vol. 46, pp. 627–643, May 1998.

[16] _____, "Cumulant-based adaptive multi-channel filtering for wireless communication systems with multipath RF propagation using antenna arrays," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 377–391, May 1998.
[17] 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.
[18] M. Moonen, P. Van Dooren, and J. Vanderwalle, "An SVD updating algorithm for subspace tracking," *SIAM J. Matrix Anal. Appl.*, vol. 13, no. 4, pp. 785–798, 1992.
[19] B. Porat, *Digital Processing of Random Signals.* Englewood Cliffs, NJ: Prentice Hall, 1994.
[20] G. W. Stewart, "A Jacobi-like algorithm for computing the Schur decomposition of a nonhermitian matrix," *SIAM J. Sci. Stat. Comp.*, vol. 6, no. 4, pp. 853–863, 1985.
[21] TIA/EIA/IS-136.1-A, "TDMA cellular/PCS—radio interface—mobile station—base station compatibility—digital control channel," and TIA/EIA/IS-136.2-A, "TDMA cellular/PCS—radio interface—mobile station—base station compatibility—traffic channels and FSK control channel," Oct. 1996.
[22] S. Van Huffel and J. Vandewalle, "The total least squares problem: Computational aspects and analysis." Philadelphia, SIAM, 1991.
[23] 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.

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

He worked for the Italian Air Force in 1990–1991 and consulted in the area of digital signal processing 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 of Alenia Spazio where he was involved in the design of satellite receivers and spread spectrum transponders. He also collaborated with the digital communications research group of Fondazione Ugo Bordoni. In 1994, he was appointed Visiting Scientist at the ECSE Department of Rensselaer Polytechnic Institute, Troy, NY. He was a wireless communications consultant for ATS Inc., Waltham, MA, and in 1995 he joined the Telecommunications Group of Watkins–Johnson Company, Gaithersburg, MD, where he is currently Head of the Advanced Wireless Development section. He has been a leader in the development of several proprietary signal processing algorithms and digital hardware platforms used in Watkins–Johnson state-of-the-art wireless products. 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.