Adaptive Interference Suppression for DS-CDMA Systems Based on Interpolated FIR Filters With Adaptive Interpolators in Multipath Channels

Rodrigo C. de Lamare and Raimundo Sampaio-Neto

***Abstract*—In this paper, we propose an adaptive linear-receiver structure based on interpolated finite-impulse response (FIR) filters with adaptive interpolators for direct-sequence code-division multiple-access (DS-CDMA) systems in multipath channels. The interpolated minimum mean-squared error (MMSE) and the interpolated constrained minimum-variance (CMV) solutions are described for a novel scheme, where the interpolator is rendered time-varying in order to mitigate multiple-access interference and multiple-path propagation effects. Based upon the interpolated MMSE and CMV solutions, we present computationally efficient stochastic gradient and exponentially weighted recursive least squares type algorithms for both receiver and interpolator filters in the supervised and blind modes of operation. A convergence analysis of the algorithms and a discussion of the convergence properties of the method are carried out for both modes of operation. Simulation experiments for a downlink scenario show that the proposed structures achieve a superior bit-error-rate convergence and steady-state performance to previously reported reduced-rank receivers at lower complexity.**

***Index Terms*—Adaptive algorithms, direct-sequence code-division multiple access (DS-CDMA), multiuser detection, reduced-rank receivers.**

I. Introduction

ADAPTIVE linear receivers [1], [3], [4], [19] are a highly effective structure for combating interference in direct-sequence code-division multiple-access (DS-CDMA) systems since they usually show good performance and have simple adaptive implementation. The linear minimum mean-squared error (MMSE) receiver [3], [4] implemented with an adaptive filter is one of the most prominent design criteria for DS-CDMA systems. Such a receiver only requires the timing of the desired user and a training sequence in order to suppress interference. Conversely, when a receiver loses track of the desired user and a training sequence is not available, a blind linear minimum-variance (MV) receiver [5], [6] that trades off the need for a training sequence in favor of the knowledge of the desired user's spreading code can be used to retrieve the desired signal.

The works in [1]–[6] were restricted to systems with short codes, where the spreading sequences are periodic. However, adaptive techniques are also applicable to systems with long codes, provided some modifications are carried out. The designer can resort to chip equalization [7] followed by a despreader for downlink scenarios. For an uplink solution, channel-estimation algorithms for aperiodic sequences [8], [9] are required, and the sample average approach for estimating the covariance matrix $\mathbf{R} = E[\mathbf{r}(i)\mathbf{r}^{\mathrm{H}}(i)]$ of the observed data $\mathbf{r}(i)$ has to be replaced by $\hat{\mathbf{R}} = \mathbf{P}\mathbf{P}^{\mathrm{H}} + \sigma^2\mathbf{I}$, which is constructed with a matrix $\mathbf{P}$ containing the effective signature sequence of users and the variance σ^2 of the receiver noise [10]. In addition, with some recent advances in random-matrix theory [11], it is also possible to deploy techniques originally developed for systems with short codes in implementations with long codes. Furthermore, the adaptive-receiver structures reported so far [1]–[10] can be easily extended to asynchronous systems in uplink connections. In the presence of large relative delays among the users, the observation window of each user should be expanded in order to consider an increased number of samples derived from the offsets among users. Alternatively, for small relative delays among users, the designer can utilize chip oversampling to compensate for the random-timing offsets. These remedies imply in augmented filter lengths and, consequently, increased computational complexity.

In this context, a problem arises when the processing gain used in the system and the number of parameters for estimation is large. In these scenarios, the receiver has to cope with difficulties such as significant computational burden, increased amount of training and poor convergence, and tracking performance. In general, when an adaptive filter with a large number of taps is used to suppress interference, then it implies slow response to changing interference and channel conditions. Reduced-rank interference suppression for DS-CDMA [12]–[18] was originally motivated by situations where the number of elements in the receiver is large and where it is desirable to work with fewer parameters for complexity and convergence reasons. Early works in reduced-rank interference suppression for DS-CDMA systems [12]–[14] were based on

Manuscript received June 27, 2005; revised March 22, 2006, July 4, 2006, February 2, 2007, and March 8, 2007. This work was supported by the Brazilian Council for Scientific and Technological Development (CNPq). The review of this paper was coordinated by Dr. M. Reed.

R. C. de Lamare was with Center for Studies in Telecommunications, Pontifical Catholic University of Rio de Janeiro, 22453-900 Rio de Janeiro, Brazil. He is now with the Communications Research Group, Department of Electronics, University of York, Y010 5DD York, U.K. (e-mail: rcdl500@ohm.york.ac.uk).

R. Sampaio-Neto is with Center for Studies in Telecommunications, Pontifical Catholic University of Rio de Janeiro, 22453-900, Rio de Janeiro, Brazil (e-mail: raimundo@cetuc.puc-rio.br).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2007.899931

principal components (PC) of the covariance matrix $\mathbf{R}$ of the observation data. This requires a computationally expensive eigendecomposition to extract the signal subspace, which leads to poor performance in systems with moderate to heavy loads. An attempt to reduce the complexity of PC approaches was reported in [15] with the partial-despreading (PD) method, where the authors report a simple technique that allows the selection of the performance between the matched filter and the full-rank MMSE receiver. A promising reduced-rank technique for interference suppression, denoted as multistage Wiener filter (MWF), was developed by Goldstein *et al.* in [17] and was later extended to stochastic gradient (SG) and recursive adaptive versions by Honig and Goldstein in [18]. A problem with the MWF approach is that, although it is less complex than the full-rank solution, it still presents a considerable computational burden and numerical problems for implementation. In this paper, we present an alternative reduced-rank interference-suppression scheme based on interpolated finite-impulse response (IFIR) filters with adaptive interpolators that gathers simplicity, great flexibility, low complexity, and high performance.

The IFIR filter is a single-rate structure that is mathematically related to signal decimation followed by filtering with a reduced number of elements [20], [21]. The basic idea is to exploit the coefficient redundancy in order to remove a number of impulse-response samples, which are recreated using an interpolation scheme. The savings are obtained by interpolating the impulse response and by decimating the interpolated signal. This technique exhibits desirable properties, such as guaranteed stability, absence of limit cycles, and low computational complexity. Thus, adaptive IFIR (AIFIR) filters [22], [23] represent an interesting alternative for substituting classical adaptive FIR filters. In some applications, they show better convergence rate and can reduce the computational burden for filtering and coefficient updating, due to the reduced number of adaptive elements. These structures have been extensively applied in the context of digital filtering, although their use for parameter estimation in communications remains unexplored.

Interference suppression with IFIR filters and time-varying interpolators with batch methods, which require matrix inversions, were reported in [24]. In this paper, we investigate the suppression of multiple-access interference (MAI) and intersymbol interference (ISI) with AIFIR filters (that do not need matrix inversions) for both supervised and blind modes of operation in synchronous DS-CDMA systems with short codes. A novel AIFIR scheme, where the interpolator is rendered adaptive, is discussed and designed with both MMSE and MV design criteria. The new scheme, introduced in [25] and [26], yields a performance superior to conventional AIFIR schemes [22], [23] (where the interpolator is fixed) and a faster convergence performance than full-rank and other existing reduced-rank receivers. Computationally efficient SG and recursive least squares (RLS)-type adaptive algorithms are developed for the new structure based upon the MMSE and MV performance criteria with appropriate constraints to mitigate MAI and ISI and jointly estimate the channel. The motivation for the novel structure is to exploit the redundancy found in DS-CDMA signals that operate in multipath, by removing a number of samples of the received signal and retrieving them through interpolation. The gains in convergence performance over full-rank solutions are obtained through the reduction of the number of parameters for estimation, leading to a faster acquisition of the required statistics of the method and a smaller misadjustment noise provided by the smaller filter [1], [19]. Furthermore, the use of an adaptive interpolator can provide a time-varying and rapid means of compensation for the decimation process and the discarded samples. The novel scheme has the potential and flexibility to consistently yield faster convergence than the full-rank approach, since the designer can choose the number of adaptive elements during the transient process and, upon convergence increase, the number of elements up to the full-rank. Unlike PC techniques, our scheme is very simple because it does not require eigendecomposition, and its performance is not severely degraded when the system load is increased. In contrast to PD, the AIFIR structure jointly optimizes two filters, namely, the interpolator and the reduced-rank, resulting in reduced-rank filters with fewer taps and faster convergence than PD since the interpolator helps with the compensation of the discarded samples. In comparison with the MWF, the proposed scheme is simpler, more flexible, and more suitable for implementation, because the MWF has numerical problems in fixed-point implementations.

A convergence analysis of the algorithms and a discussion of the global convergence properties of the method, which are not treated in [24]–[26], are undertaken for both modes of operation. Specifically, we study the convergence properties of the proposed joint adaptive interpolator and receiver scheme and conclude that it leads to an optimization problem with multiple global minima and no local minima. In this regard and based on the analyzed convergence properties of the method, we show that the prediction of the excess MSE of both blind and supervised adaptive algorithms is rendered possible through the study of the MSE trajectory of only one of the jointly optimized parameter vectors, i.e., the interpolator or the reduced-rank filters. Then, using common assumptions of the adaptive filtering literature, such as the independence theory, we analyze the trajectory of the mean tap vector of the joint optimization of the interpolator and the receiver and MSE trajectory. We also provide some mathematical conditions, which explain why the new scheme with SG- and RLS-type algorithms is able to converge faster than the full-rank scheme. Although the novel structure and algorithms are examined in a synchronous downlink scenario with periodic signature sequences in this paper, it should be remarked that they can be extended to long codes and asynchronous systems, provided that the designer adopts the modifications explained in the works reported in [7]–[11].

This paper is organized as follows. Section II describes the DS-CDMA system model. The linear-interpolated-receiver principle and design criteria, namely, the MMSE and constrained MV (CMV), are described in Section III. Section IV is dedicated to the derivation of adaptive algorithms, and Section V is devoted to the global convergence properties of the method and the convergence analysis of the algorithms. Section VI presents and discusses the simulation results, and Section VII gives the concluding remarks.

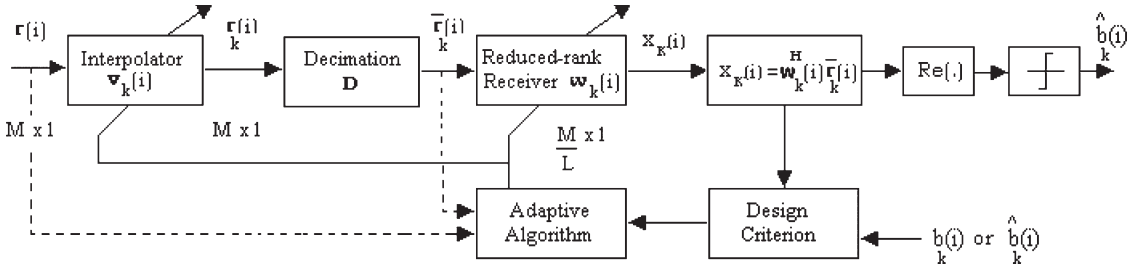

Fig. 1. Proposed adaptive reduced-rank receiver structure.

II. DS-CDMA System Model

Let us consider the downlink of a synchronous DS-CDMA system with K users, N chips per symbol, and L_p propagation paths. The signal broadcasted by the base station intended for user k has a baseband representation given by

$$x_k(t) = A_k \sum_{i=-\infty}^{\infty} b_k(i) s_k(t - iT) \tag{1}$$

where $b_k(i) \in \{\pm 1\}$ denotes the ith symbol for user k, the real-valued spreading waveform, and the amplitude associated with user k are $s_k(t)$ and A_k, respectively. The spreading waveforms are expressed by $s_k(t) = \sum_{i=1}^{N} a_k(i)\phi(t - iT_c)$, where $a_k(i) \in \{\pm 1/\sqrt{N}\}$, $\phi(t)$ is the chip waveform, T_c is the chip duration, and $N = T/T_c$ is the processing gain. Assuming that the receiver is synchronized with the main path, the coherently demodulated composite received signal is

$$r(t) = \sum_{k=1}^{K} \sum_{l=0}^{L_p - 1} h_l(t) x_k(t - \tau_l) + n(t) \tag{2}$$

where $h_l(t)$ and τ_l are, respectively, the channel coefficient and the delay associated with the lth path. Assuming that $\tau_{k,l} = lT_c$, the channel is constant during each symbol interval, and the spreading codes are repeated from symbol to symbol, the received signal $r(t)$ after filtering by a chip-pulse matched filter and sampled at chip rate yields the $M = N + L_p - 1$ dimensional received vector

$$\mathbf{r}(i) = \mathbf{H}(i) \sum_{k=1}^{K} A_k \mathbf{S}_k \mathbf{b}_k(i) + \mathbf{n}(i) \tag{3}$$

where $\mathbf{n}(i) = [n_1(i) \cdots n_M i)]^{\mathrm{T}}$ is the complex Gaussian noise vector with $E[\mathbf{n}(i)\mathbf{n}^{\mathrm{H}}(i)] = \sigma^2 \mathbf{I}$, where $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denotes transpose and Hermitian transpose, respectively, and $E[.]$ is the expected value; the kth user symbol vector is $\mathbf{b}_k(i) = [b_k(i + L_s - 1) \cdots b_k(i) \cdots b_k(i - L_s + 1)]^{\mathrm{T}}$, where L_s is the ISI span, and the $((2L_s - 1) \times N) \times (2L_s - 1)$ matrix $\mathbf{S}_k$ with nonoverlapping shifted versions of the signature of user k is

$$\mathbf{S}_k = \begin{bmatrix} \mathbf{s}_k & 0 & \cdots & 0 \\ 0 & \mathbf{s}_k & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \mathbf{s}_k \end{bmatrix} \tag{4}$$

where the signature sequence for the kth user is $\mathbf{s}_k = [a_k(1) \cdots a_k(N)]^{\mathrm{T}}$, and the $M \times ((2L_s - 1) \times N)$ channel matrix $\mathbf{H}(i)$ is

$$\mathbf{H}(i) = \begin{bmatrix} h_0(i) & \cdots & h_{L_p-1}(i) & \cdots & 0 & 0 \\ 0 & h_0(i) & \cdots & h_{L_p-1}(i) & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & h_0(i) & \cdots & h_{L_p-1}(i) \end{bmatrix} \tag{5}$$

where $h_l(i) = h_l(iT_c)$ for $l = 0, \ldots, L_p - 1$. The MAI arises from the nonorthogonality between the received signals, whereas the ISI span L_s depends on the length of the channel response, which is related to the length of the chip sequence. For $L_p = 1$, $L_s = 1$ (no ISI), for $1 < L_p \leqslant N$, $L_s = 2$, for $N < L_p \leqslant 2N$, $L_s = 3$, and so on.

III. Linear Interpolated CDMA Receivers

The underlying principles of the proposed CDMA-receiver structure are detailed here. Fig. 1 shows the structure of an IFIR receiver, where an interpolator and a reduced-rank receiver that are time-varying are employed. The $M \times 1$ received vector $\mathbf{r}(i) = [r_0^{(i)} \cdots r_{M-1}^{(i)}]^{\mathrm{T}}$ is filtered by the interpolator filter $\mathbf{v}_k(i) = [v_{k,0}^{(i)} \cdots v_{k,N_I-1}^{(i)}]^{\mathrm{T}}$, yielding the interpolated received vector $\mathbf{r}_k(i)$. The vector $\mathbf{r}_k(i)$ is then projected onto an $M/L \times 1$ vector $\bar{\mathbf{r}}_k(i)$. This procedure corresponds to removing $L - 1$ samples of $\mathbf{r}_k(i)$ of each set of L consecutive ones. Then, the inner product of $\bar{\mathbf{r}}_k(i)$ with the M/L-dimensional vector of filter coefficients $\mathbf{w}_k(i) = [w_{k,0}^{(i)} \cdots w_{k,M/L-1}^{(i)}]^{\mathrm{T}}$ is computed to yield the output $x_k(i)$.

The projected interpolated observation vector $\bar{\mathbf{r}}_k(i) = \mathbf{D}\mathbf{r}_k(i)$ is obtained with the aid of the $M/L \times M$ projection matrix $\mathbf{D}$, which is mathematically equivalent to signal decimation on the $M \times 1$ vector $\mathbf{r}_k(i)$. An interpolated receiver with decimation factor L can be designed by choosing $\mathbf{D}$ as

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \underbrace{0 & \cdots & 0}_{(m-1)L \text{ zeros}} & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \underbrace{0 & 0 & 0 & 0 & 0 & \cdots & 0}_{(M/L-1)L \text{ zeros}} & 1 & \underbrace{0 & \ldots & 0}_{(L-1) \text{ zeros}} \end{bmatrix} \tag{6}$$

where $m(m = 1, 2, \ldots, M/L)$ denotes the mth row. The strategy that allows us to devise solutions for both interpolator and receiver is to express the estimated symbol $x_k(i) = \mathbf{w}_k^{\mathrm{H}}(i)\bar{\mathbf{r}}_k(i)$ as a function of $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$ [we will drop the subscript k and symbol index (i) for ease of presentation]

$$\begin{aligned} x_k(i) &= w_0^* \mathbf{v}_k^{\mathrm{H}} \dot{\mathbf{r}}_0 + w_1^* \mathbf{v}_k^{\mathrm{H}} \dot{\mathbf{r}}_1 + \cdots + w_{M/L-1}^* \mathbf{v}_k^{\mathrm{H}} \dot{\mathbf{r}}_{M/L-1} \\ &= \mathbf{v}_k^{\mathrm{H}}(i) \left[\dot{\mathbf{r}}_0^{(i)} | \cdots | \dot{\mathbf{r}}_{M/L-1}^{(i)} \right] \mathbf{w}_k^*(i) \\ &= \mathbf{v}_k^{\mathrm{H}}(i) \mathfrak{R}(i) \mathbf{w}_k^*(i) \end{aligned} \tag{7}$$

where $\mathbf{u}_k(i) = \mathfrak{R}(i)\mathbf{w}_k^*(i)$ is an $N_I \times 1$ vector, the M/L coefficients of $\mathbf{w}_k(i)$ and the N_I elements of $\mathbf{v}_k(i)$ are assumed to be complex, the asterisk denotes complex conjugation, and $\dot{\mathbf{r}}_s(i)$ is a length N_I segment of the received vector $\mathbf{r}(i)$ beginning at $r_{s\times L}(i)$, and

$$\mathfrak{R}(i) = \begin{bmatrix} r_0^{(i)} & r_L^{(i)} & \cdots & r_{(M/L-1)L}^{(i)} \\ r_1^{(i)} & r_{L+1}^{(i)} & \cdots & r_{(M/L-1)L+1}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N_I-1}^{(i)} & r_{L+N_I-1}^{(i)} & \cdots & r_{(M/L-1)L+N_I-1}^{(i)} \end{bmatrix}. \tag{8}$$

The interpolated linear-receiver design is equivalent in determining an FIR filter $\mathbf{w}_k(i)$ with M/L coefficients that provide an estimate of the desired symbol

$$\hat{b}_k(i) = \mathrm{sgn}\left(\mathrm{Re}\left[\mathbf{w}_k^{\mathrm{H}}(i)\bar{\mathbf{r}}_k(i)\right]\right) \tag{9}$$

where $\mathrm{Re}(\cdot)$ selects the real part, $\mathrm{sgn}(\cdot)$ is the signum function, and the receiver parameter vector $\mathbf{w}_k$ is optimized according to a selected design criterion.

A. MMSE Reduced-Rank Interpolated Receiver Design

The MMSE solutions for $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$ can be computed if we consider the optimization problem whose cost function is

$$J_{\mathrm{MSE}}\left(\mathbf{w}_k(i), \mathbf{v}_k(i)\right) = E\left[|b_k(i) - \mathbf{v}_k^{\mathrm{H}}(i)\mathfrak{R}(i)\mathbf{w}_k^*(i)|^2\right] \tag{10}$$

where $b_k(i)$ is the desired symbol for user k at time index (i). By fixing the interpolator $\mathbf{v}_k(i)$ and minimizing (10) with respect to $\mathbf{w}_k(i)$, the interpolated Wiener filter/receiver weight vector is

$$\mathbf{w}_k(i) = \boldsymbol{\alpha}(\mathbf{v}_k) = \bar{\mathbf{R}}_k^{-1}(i)\bar{\mathbf{p}}_k(i) \tag{11}$$

where $\bar{\mathbf{R}}_k(i) = E[\bar{\mathbf{r}}_k(i)\bar{\mathbf{r}}_k^{\mathrm{H}}(i)]$, $\bar{\mathbf{p}}_k(i) = E[b_k^*(i)\bar{\mathbf{r}}_k(i)]$, and $\bar{\mathbf{r}}_k(i) = \mathfrak{R}^{\mathrm{T}}(i)\mathbf{v}_k^*(i)$, and by fixing $\mathbf{w}_k(i)$ and minimizing (10) with respect to $\mathbf{v}_k(i)$, the interpolator weight vector is

$$\mathbf{v}_k(i) = \boldsymbol{\beta}(\mathbf{w}_k) = \mathbf{R}_{u_k}^{-1}(i)\mathbf{p}_{u_k}(i) \tag{12}$$

where $\mathbf{R}_{u_k}(i) = E[\mathbf{u}_k(i)\mathbf{u}_k^{\mathrm{H}}(i)]$, $\mathbf{p}_{u_k}(i) = E[b_k^*(i)\mathbf{u}_k(i)]$, and $\mathbf{u}_k(i) = \mathfrak{R}(i)\mathbf{w}_k^*(i)$. The associated MSE expressions are

$$\begin{aligned} J(\mathbf{v}_k) &= J_{\mathrm{MSE}}\left(\boldsymbol{\alpha}(\mathbf{v}_k), \mathbf{v}_k\right) \\ &= \sigma_b^2 - \bar{\mathbf{p}}_k^{\mathrm{H}}(i)\mathbf{R}_k^{-1}(i)\bar{\mathbf{p}}_k(i) \end{aligned} \tag{13}$$

$$J_{\mathrm{MSE}}\left(\mathbf{w}_k, \boldsymbol{\beta}(\mathbf{w}_k)\right) = \sigma_b^2 - \mathbf{p}_{u_k}^{\mathrm{H}}(i)\mathbf{R}_{u_k}^{-1}(i)\mathbf{p}_{u_k}(i) \tag{14}$$

where $\sigma_b^2 = E[|b(i)|^2]$. Note that points of global minimum of (10) can be obtained by $\mathbf{v}_{k,\mathrm{opt}} = \arg\min_{\mathbf{v}_k} J(\mathbf{v}_k)$ and $\mathbf{w}_{k,\mathrm{opt}} = \boldsymbol{\alpha}(\mathbf{v}_{k,\mathrm{opt}})$ or $\mathbf{w}_{k,\mathrm{opt}} = \arg\min_{\mathbf{w}_k} J_{\mathrm{MSE}}(\mathbf{w}_k, \boldsymbol{\beta}(\mathbf{w}_k))$ and $\mathbf{v}_{k,\mathrm{opt}} = \boldsymbol{\beta}(\mathbf{w}_{k,\mathrm{opt}})$. At the minimum point, (13) equals (14), and the MMSE for the proposed structure is achieved. We remark that (11) and (12) are not closed-form solutions for $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$, since (11) is a function of $\mathbf{v}_k(i)$, and (12) depends on $\mathbf{w}_k(i)$, and thus, it is necessary to iterate (11) and (12) with an initial guess to obtain a solution, as in [24]. An iterative MMSE solution can be sought via adaptive algorithms.

B. CMV Reduced-Rank Interpolated Receiver Design

The interpolated CMV receiver parameter vector $\mathbf{w}_k$ and the interpolator parameter vector $\mathbf{v}_k$ are obtained by minimizing

$$\begin{aligned} J_{\mathrm{MV}}(\mathbf{w}_k, \mathbf{v}_k) &= E\left[|x_k(i)|^2\right] = E\left[\left|\mathbf{v}_k^{\mathrm{H}}(i)\mathfrak{R}(i)\mathbf{w}_k^*(i)\right|^2\right] \\ &= \mathbf{w}_k^{\mathrm{H}}(i)\bar{\mathbf{R}}_k\mathbf{w}_k(i) = \mathbf{v}_k^{\mathrm{H}}(i)\mathbf{R}_{\mathbf{u_k}}\mathbf{v}_k(i) \end{aligned} \tag{15}$$

subject to the proposed constraints $\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{w}_k(i) = \mathbf{g}(i)$ and $\|\mathbf{v}_k(i)\| = 1$, where the $M \times L_p$ constraint matrix $\mathbf{C}_k$ contains one-chip shifted versions of the signature sequence of user k; $\mathbf{g}(i)$ is an L_p-dimensional parameter vector to be determined. The vector of constraints $\mathbf{g}(i)$ can be chosen among various criteria, although, in this paper, we adopt $\mathbf{g}(i)$ as the channel parameter vector ($\mathbf{g} = [h_0 \cdots h_{L_p-1}]^{\mathrm{T}}$) because it provides better performance than other choices, as reported in [6]. The proposed constraint $\|\mathbf{v}_k(i)\| = 1$ ensures adequate design values for the interpolator filter $\mathbf{v}_k$, whereas $\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{w}_k(i) = \mathbf{g}(i)$ avoids the suppression of the desired signal. By fixing $\mathbf{v}_k$, taking the gradient of the Lagrangian function $J_{\mathrm{MV}}^l(\mathbf{w}_k, \mathbf{v}_k) = E[[|\mathbf{v}_k^{\mathrm{H}}(i)\mathfrak{R}(i)\mathbf{w}_k^*(i)|^2] + \mathrm{Re}[(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{w}_k(i) - \mathbf{g}(i))^{\mathrm{H}}\boldsymbol{\lambda}]$, where $\boldsymbol{\lambda}$ is a vector of Lagrange multipliers, with respect to $\mathbf{w}_k$ and setting it to $\mathbf{0}$, we get

$$\begin{aligned} E\left[\bar{\mathbf{r}}_k(i)\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\right]\mathbf{w}_k(i) + \mathbf{D}\mathbf{C}_k\boldsymbol{\lambda} = \mathbf{0} \\ \Longrightarrow \mathbf{w}_k(i) = -\bar{\mathbf{R}}_k^{-1}(i)\mathbf{D}\mathbf{C}_k\boldsymbol{\lambda}. \end{aligned}$$

Using the constraint $\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{w}_k(i) = \mathbf{g}(i)$ and substituting $\mathbf{w}_k(i) = -\bar{\mathbf{R}}_k^{-1}(i)\mathbf{D}\mathbf{C}_k\boldsymbol{\lambda}$, we arrive at $\boldsymbol{\lambda} = -(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\bar{\mathbf{R}}_k^{-1}(i)\mathbf{D}\mathbf{C}_k)^{-1}\mathbf{g}_k(i)$. The resulting expression for the receiver is

$$\begin{aligned} \mathbf{w}_k(i) &= \boldsymbol{\alpha}_o(\mathbf{v}_k) \\ &= \bar{\mathbf{R}}_k(i)^{-1}\mathbf{D}\mathbf{C}_k\left(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\bar{\mathbf{R}}_k(i)^{-1}\mathbf{D}\mathbf{C}_k\right)^{-1}\mathbf{g}(i) \end{aligned} \tag{16}$$

and the associate minimum output variance is

$$\begin{aligned} J_o(\mathbf{v}_k) &= J_{\mathrm{MV}}\left(\boldsymbol{\alpha}_o(\mathbf{v}_k), \mathbf{v}_k\right) = \mathbf{w}_k^{\mathrm{H}}(i)\bar{\mathbf{R}}_k(i)\mathbf{w}_k(i) \\ &= \mathbf{g}^{\mathrm{H}}(i)\left(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\bar{\mathbf{R}}_k(i)^{-1}\mathbf{D}\mathbf{C}_k\right)^{-1}\mathbf{g}(i). \end{aligned} \tag{17}$$

By fixing $\mathbf{w}_k$, the solution that minimizes (15) is

$$\mathbf{v}_k(i) = \boldsymbol{\beta}_o(\mathbf{w}_k) = \arg\min_{\mathbf{v}} \mathbf{v}^{\mathrm{H}}\mathbf{R}_{\mathbf{u}_k}(i)\mathbf{v} \tag{18}$$

subject to $\|\mathbf{v}_k(i)\| = 1$. Therefore, the solution for the interpolator is the normalized eigenvector of $\mathbf{R}_{\mathbf{u_k}}$ corresponding to its minimum eigenvalue, via singular-value decomposition (SVD). As occurring with the MMSE approach, we iterate (16) and (18) with an initial guess to obtain a CMV solution [24]. Note also that (16) assumes the knowledge of the channel parameters. However, in applications where multipath is present, these parameters are not known, and thus, channel estimation is required. To blindly estimate the channel, we use the method of [6], [27]

$$\hat{\mathbf{g}}(i) = \arg\min_{\mathbf{g}} \mathbf{g}^{\mathrm{H}}\mathbf{C}_k^{\mathrm{H}}\mathbf{R}^{-m}(i)\mathbf{C}_k\mathbf{g} \tag{19}$$

subject to $\|\hat{\mathbf{g}}\| = 1$, where $\mathbf{R}(i) = E[\mathbf{r}(i)\mathbf{r}^{\mathrm{H}}(i)]$, and m is a finite power and whose solution is the eigenvector corresponding to the minimum eigenvalue of the $L_p \times L_p$ matrix $\mathbf{C}_k^T\mathbf{R}(i)^{-m}\mathbf{C}_k$ through SVD. Note that, in this paper, we restrict the values of m to 1, although the performance of the channel estimator and, consequently, of the receiver can be improved by increasing m. In the next section, we propose iterative solutions via adaptive algorithms.

IV. Adaptive Algorithms

We describe SG and RLS algorithms [31, Ch. 9 and 13] that adjust the parameters of the receiver and the interpolator based on the MMSE criterion and the constrained minimization of the MV cost function [25], [26]. The novel structure, as shown in Fig. 1 and denoted as INT, for the receivers gathers fast convergence, low complexity, and additional flexibility, since the designer can adjust the decimation factor L and the length of the interpolator N_I, depending on the needs of the application and the hostility of the environment. Based upon the MMSE and CMV design criteria, the proposed receiver structure has the following modes of operation: training mode, where it employs a known training sequence; decision directed mode, which uses past decisions in order to estimate the receiver parameters; and blind mode, which is based on the CMV criterion and trades off the training sequence against the knowledge of the signature sequence. The complexity, in terms of arithmetic operations of the algorithms associated with the INT and the existing techniques, is included as a function of the number of adaptive elements for comparison purposes.

A. Least Mean Squares (LMS) Algorithm

Given the projected interpolated observation vector $\bar{\mathbf{r}}_k(i)$ and the desired symbol $b_k(i)$, we consider the following cost function:

$$J_{\mathrm{MSE}} = |b_k(i) - \mathbf{v}_k^{\mathrm{H}}(i)\Re(i)\mathbf{w}_k^*(i)|^2. \tag{20}$$

Taking the gradient terms of (20) with respect to $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$ and using the gradient descent rules [31, Ch. 9, pp. 367–371], $\mathbf{w}_k(i+1) = \mathbf{w}_k(i) - \mu\nabla_{\mathbf{w}^*}^{J_{\mathrm{MSE}}}$ and $\mathbf{v}_k(i+1) = \mathbf{v}_k(i) - \eta\nabla_{\mathbf{v}^*}^{J_{\mathrm{MSE}}}$ yields

$$\mathbf{v}_k(i+1) = \mathbf{v}_k(i) + \eta e_k^*(i)\mathbf{u}_k(i) \tag{21}$$

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) + \mu e_k^*(i)\bar{\mathbf{r}}_k(i) \tag{22}$$

where $e_k(i) = b_k(i) - \mathbf{w}_k(i)^{\mathrm{H}}\bar{\mathbf{r}}_k(i)$ is the error for user k, $\mathbf{u}_k = \Re(i)\mathbf{w}_k(i)$, and μ and η are the step sizes of the algorithm for $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$. The LMS algorithm for the proposed structure described in this section has a computational complexity $O(M/L + N_I)$. In fact, the proposed structure trades off one LMS algorithm with complexity $O(M)$ against two LMS algorithms with complexity $O(M/L)$ and $O(N_I)$, operating in parallel. It is worth noting that, to stabilize and to facilitate tuning of parameters, it is useful to employ normalized step sizes and, consequently, normalized least mean squares (NLMS)-type recursions when operating in a changing environment, and thus, we have $\mu(i) = \eta_0/\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\bar{\mathbf{r}}_k(i)$ and $\eta(i) = \mu_0/\mathbf{u}_k^{\mathrm{H}}(i)\mathbf{u}_k(i)$ as the step sizes of the algorithm for $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$, where μ_0 and η_0 are the convergence factors.

B. RLS Algorithm

Consider the time average estimate of the matrix $\bar{\mathbf{R}}_k$, as required in (11), given by $\hat{\bar{\mathbf{R}}}_k(i) = \sum_{l=1}^{i}\alpha^{i-l}\bar{\mathbf{r}}_k(l)\bar{\mathbf{r}}_k^{\mathrm{H}}(l)$, where α $(0 < \alpha \leqslant 1)$ is the forgetting factor that can be alternatively expressed by $\hat{\bar{\mathbf{R}}}_k(i) = \alpha\hat{\bar{\mathbf{R}}}_k(i-1) + \bar{\mathbf{r}}_k(i)\bar{\mathbf{r}}_k^{\mathrm{H}}(i)$. To avoid the inversion of $\hat{\bar{\mathbf{R}}}_k(i)$ required in (11), we use the matrix-inversion lemma and define $\mathbf{P}_k(i) = \hat{\bar{\mathbf{R}}}_k^{-1}(i)$ and the gain vector $\mathbf{G}_k(i)$ as

$$\mathbf{G}_k(i) = \frac{\alpha^{-1}\mathbf{P}_k(i-1)\bar{\mathbf{r}}_k(i)}{1 + \alpha^{-1}\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\mathbf{P}_k(i-1)\bar{\mathbf{r}}_k(i)} \tag{23}$$

and thus, we can rewrite $\mathbf{P}_k(i)$ as

$$\mathbf{P}_k(i) = \alpha^{-1}\mathbf{P}_k(i-1) - \alpha^{-1}\mathbf{G}_k(i)\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\mathbf{P}_k(i-1). \tag{24}$$

By rearranging (23), we have $\mathbf{G}_k(i) = \alpha^{-1}\mathbf{P}_k(i-1)\bar{\mathbf{r}}_k(i) - \alpha^{-1}\ \mathbf{G}_k(i)\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\mathbf{P}_k(i-1)\bar{\mathbf{r}}_k(i) = \mathbf{P}_k(i)\bar{\mathbf{r}}_k(i)$. By employing the least squares (LS) solution [a time average of (11)] and the recursion $\hat{\mathbf{p}}_k(i) = \alpha\hat{\mathbf{p}}_k(i-1) + \bar{\mathbf{r}}_k(i)b_k^*(i)$, we obtain

$$\mathbf{w}_k(i) = \hat{\bar{\mathbf{R}}}_k^{-1}(i)\hat{\mathbf{p}}_k(i) = \alpha\mathbf{P}_k(i)\hat{\mathbf{p}}_k(i-1) + \mathbf{P}_k(i)\bar{\mathbf{r}}_k(i)b_k^*(i). \tag{25}$$

Substituting (24) into (25) yields

$$\mathbf{w}_k(i) = \mathbf{w}_k(i-1) + \mathbf{G}_k(i)\xi_k^*(i) \tag{26}$$

where the *a priori* estimation error is described by $\xi_k(i) = b_k(i) - \mathbf{w}_k^{\mathrm{H}}(i-1)\bar{\mathbf{r}}_k(i)$. Similar recursions for the interpolator are devised by using (12). The estimate $\hat{\mathbf{R}}_{\mathbf{u}_k}$ can be obtained through $\hat{\mathbf{R}}_{\mathbf{u}_k}(i) = \sum_{l=1}^{i}\alpha^{i-l}\mathbf{u}_k(l)\mathbf{u}_k^{\mathrm{H}}(l)$ and can be alternatively written as $\hat{\mathbf{R}}_{\mathbf{u}_k}(i) = \alpha\hat{\mathbf{R}}_{\mathbf{u}_k}(i-1) + \mathbf{u}_k(i)\mathbf{u}_k^{\mathrm{H}}(i)$. To avoid the inversion of $\hat{\mathbf{R}}_{\mathbf{u}_k}$, we use the matrix-inversion lemma, and, again, for convenience of computation, we define $\mathbf{P}_{\mathbf{u}_k}(i) = \hat{\mathbf{R}}_{\mathbf{u}_k}^{-1}(i)$ and the Kalman gain vector $\mathbf{G}_{\mathbf{u}_k}(i)$ as

$$\mathbf{G}_{\mathbf{u}_k}(i) = \frac{\alpha^{-1}\mathbf{P}_{\mathbf{u}_k}(i-1)\mathbf{u}_k(i)}{1 + \alpha^{-1}\mathbf{u}_k^{\mathrm{H}}(i)\mathbf{P}_{\mathbf{u}_k}(i-1)\mathbf{u}_k(i)} \tag{27}$$

and thus, we can rewrite (27) as

$$\mathbf{P}_{\mathbf{u}_k}(i) = \alpha^{-1}\mathbf{P}_{\mathbf{u}_k}(i-1) - \alpha^{-1}\mathbf{G}_{\mathbf{u}_k}(i)\mathbf{u}_k^{\mathrm{H}}(i)\mathbf{P}_{\mathbf{u}_k}(i-1). \tag{28}$$

By proceeding in a similar approach to the one taken to obtain (26), we arrive at

$$\mathbf{v}_k(i) = \mathbf{v}_k(i-1) + \mathbf{G}_{\mathbf{u}_k}(i)\xi_k^*(i). \tag{29}$$

The RLS algorithm for the proposed structure trades off a computational complexity of $O(M^2)$ against two RLS algorithms operating in parallel, with complexity $O((M/L)^2)$ and $O(N_I^2)$, respectively. Because N_I is small ($N_I \ll M$, as will be shown later), the computational advantage of the RLS combined with the INT structure is rather significant.

C. CMV-SG Algorithm

Consider the unconstrained Lagrangian MV cost function

$$J_{\mathrm{MV}} = \left(\mathbf{v}_k^{\mathrm{H}}(i)\mathbf{u}_k(i)\mathbf{u}_k^{\mathrm{H}}(i)\mathbf{v}_k(i)\right) + \boldsymbol{\lambda}^{\mathrm{H}}\left(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{w}_k(i) - \mathbf{g}(i)\right) + \left(\mathbf{w}_k^{\mathrm{H}}(i)\mathbf{D}\mathbf{C}_k - \mathbf{g}^{\mathrm{H}}(i)\right)\boldsymbol{\lambda} \tag{30}$$

where $\boldsymbol{\lambda}$ is a vector of Lagrange multipliers. An SG solution can be devised by taking the gradient terms of (30) with respect to $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$, as described by $\mathbf{w}_k(i+1) = \mathbf{w}_k(i) - \mu_w(i)\nabla J_{\mathbf{w}_k(i)}$ and $\mathbf{v}_k(i+1) = \mathbf{v}_k(i) - \eta(i)\nabla J_{\mathbf{v}_k(i)}$, which adaptively minimizes J_{MV} with respect to $\mathbf{w}_k(i)$ and $\mathbf{v}_k(i)$. Substituting the gradient terms, the equations become

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) - \mu_w(i)\left(x_k^*(i)\bar{\mathbf{r}}_k(i) + \mathbf{D}\mathbf{C}_k\boldsymbol{\lambda}(i)\right) \tag{31}$$

$$\mathbf{v}_k(i+1) = \mathbf{v}_k(i) - \eta(i)x_k^*(i)\mathbf{u}_k(i) \tag{32}$$

where $x_k(i) = \mathbf{w}_k^{\mathrm{H}}(i)\bar{\mathbf{r}}_k(i) = \mathbf{v}_k^{\mathrm{H}}(i)\mathbf{u}_k(i)$. We use (32) and can make $\mathbf{v}_k(i+1) \leftarrow \mathbf{v}_k(i+1)/\|\mathbf{v}_k(i+1)\|$ to update the interpolator $\mathbf{v}_k$. It is worth noting that, in our studies, the normalization on SG algorithms does not lead to different results from the ones obtained with a nonnormalized interpolator recursion. In this regard, analyzing the convergence of (32) without normalization is mathematically simpler and gives us the necessary insight into its convergence. By combining the constraint $\mathbf{C}_k\mathbf{D}^{\mathrm{H}}\mathbf{w}_k(i) = \mathbf{g}(i)$ and (32), we obtain the Lagrange multiplier

$$\boldsymbol{\lambda}(i) = (\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{D}\mathbf{C}_k)^{-1} \times \left(\mathbf{C}^{\mathrm{H}}\mathbf{D}\mathbf{w}_k(i) - \mu_w\mathbf{C}^{\mathrm{H}}\mathbf{D}x_k^*(i)\bar{\mathbf{r}}_k(i) - \mathbf{g}(i)\right). \tag{33}$$

By substituting (33) into (31), we arrive at the update rules for the estimation of the parameters of the receiver $\mathbf{w}_k$

$$\mathbf{w}_k(i+1) = \boldsymbol{\Pi}_k\left(\mathbf{w}_k(i) - \mu_w(i)x_k^*(i)\bar{\mathbf{r}}_k(i)\right) + \mathbf{D}\mathbf{C}_k\left(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{D}\mathbf{C}_k\right)^{-1}\mathbf{g}(i) \tag{34}$$

where $\boldsymbol{\Pi}_k = \mathbf{I} - \mathbf{D}\mathbf{C}_k(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{D}\mathbf{C}_k)^{-1}\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}$ is a matrix that projects $\mathbf{w}_k$ onto another hyperplane to ensure the constraints.

Normalized versions of these algorithms can be devised by substituting (32) and (34) into the MV cost function, differentiating the cost function with respect to $\mu_w(i)$ and $\mu_v(i)$, setting them to zero, and solving the new equations. Hence, the CMV-SG algorithm proposed here for the INT receiver structure adopts the normalized step sizes $\mu_w(i) = \mu_0/\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\boldsymbol{\Pi}_k\bar{\mathbf{r}}_k(i)$ and $\eta(i) = \eta_0/\mathbf{u}_k^{\mathrm{H}}\mathbf{u}_k(i)$, where μ_0 and η_0 are the convergence factors for $\mathbf{w}_k$ and $\mathbf{v}_k$, respectively.

The channel estimate $\hat{\mathbf{g}}(i)$ is obtained through the power method and the SG technique described in [27]. The method is an SG adaptive version of the blind channel-estimation algorithm described in (19) and introduced in [28] that requires only $O(L_p^2)$ arithmetic operations to estimate the channel against $O(L_p^3)$ of its SVD version. In terms of computational complexity, for the rejection of MAI and ISI, the proposed blind interpolated receiver trades off one blind algorithm with complexity $O(M)$ against two blind algorithms with complexity $O(M/L)$ and $O(N_I)$ operating in parallel.

D. CMV-RLS Algorithm

Based upon the expressions for the receiver $\mathbf{w}_k$ and interpolator $\mathbf{v}_k$ in (16) and (18) of the interpolated CMV receiver, we develop a computationally efficient RLS algorithm for the INT structure that estimates the parameters of $\mathbf{w}_k$ and $\mathbf{v}_k$.

The iterative power method [29, pp. 405–408], [30, pp. 314–333] is used in numerical analysis to compute the eigenvector corresponding to the largest singular value of a matrix. In order to obtain an estimate of $\mathbf{v}_k$ and avoid the SVD on the estimate of the matrix $\mathbf{R}_{\mathbf{u}_k}(i)$, we resort to a variation of the iterative power method to obtain the eigenvector of $\mathbf{R}_{\mathbf{u}_k}(i)$, which corresponds to the minimum eigenvalue.

Specifically, we apply the power method to the difference between $\mathbf{R}_{\mathbf{u}_k}(i)$ and the identity matrix $\mathbf{I}$ rather than applying it to the inverse of $\mathbf{R}_{\mathbf{u}_k}(i)$. This approach, which is known as shift iterations [30, p. 319], leads to computational savings on one order of magnitude, since direct SVD requires $O(N_I^3)$, while our approach needs $O(N_I^2)$. The simulations carried out reveal that this method exhibits no performance loss. Hence, we estimate $\mathbf{R}_{\mathbf{u}_k}(i)$ via the recursion $\hat{\mathbf{R}}_{\mathbf{u}_k}(i) = \sum_{n=0}^{i}\alpha^{i-n}\mathbf{u}_k(n)\mathbf{u}_k^{\mathrm{H}}(n)$ and then obtain the interpolator $\hat{\mathbf{v}}_k$ with a one-step iteration given by

$$\hat{\mathbf{v}}_k(i) = \left(\mathbf{I} - \nu_k(i)\hat{\mathbf{R}}_{\mathbf{u}_k}(i)\right)\hat{\mathbf{v}}_k(i-1) \tag{35}$$

where $\nu_k(i) = 1/\mathrm{tr}[\hat{\mathbf{R}}_{\mathbf{u}_k}(i)]$. After that, we make $\hat{\mathbf{v}}_k(i) \leftarrow \hat{\mathbf{v}}_k(i)/\|\hat{\mathbf{v}}_k(i)\|$ normalize the interpolator. This procedure is based on the following result.

Lemma 1: Let $\mathbf{R}$ be a positive semidefinite Hermitian symmetric matrix and $\mathbf{q}_{\min}$ be the eigenvector associated with the smallest eigenvalue. If $\mathbf{q}_{\min}$ is unique and of unit norm, then, with $\nu = 1/\mathrm{tr}[\mathbf{R}]$, the sequence of vectors $\mathbf{v}(i) = \hat{\mathbf{v}}(i)/\|\hat{\mathbf{v}}(i)\|$ with $\hat{\mathbf{v}}(i) = (\mathbf{I} - \nu(i)\mathbf{R})\hat{\mathbf{v}}(i-1)$ converges to $\mathbf{q}_{\min}$, provided that $\hat{\mathbf{v}}(0)$ is not orthogonal to $\mathbf{q}_{\min}$. A proof is given in the Appendix.

To recursively estimate the matrix $\bar{\mathbf{R}}_k(i)$ and avoid its inversion, we use the matrix-inversion lemma and Kalman RLS recursions [31]

$$\mathbf{G}(i) = \frac{\alpha^{-1}\hat{\bar{\mathbf{R}}}_k^{-1}(i-1)\bar{\mathbf{r}}_k(i)}{1+\alpha^{-1}\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\hat{\bar{\mathbf{R}}}_k^{-1}(i-1)\bar{\mathbf{r}}_k(i)} \quad (36)$$

$$\hat{\bar{\mathbf{R}}}_k^{-1}(i) = \alpha^{-1}\hat{\bar{\mathbf{R}}}_k^{-1}(i-1) - \alpha^{-1}\mathbf{G}(i)\bar{\mathbf{r}}_k^{\mathrm{H}}(i)\hat{\bar{\mathbf{R}}}_k^{-1}(i-1) \quad (37)$$

where $0 < \alpha \leqslant 1$ is the forgetting factor. The algorithm can be initialized with $\bar{\mathbf{R}}_k^{-1}(0) = \delta\mathbf{I}$ and $\mathbf{R}_{\mathbf{u}_k}^{-1}(0) = \delta\mathbf{I}$, where δ is a large positive number. For the computation of the reduced-rank receiver parameter vector $\mathbf{w}_k$, we use the matrix-inversion lemma [31] to estimate $(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\bar{\mathbf{R}}_k^{-1}(i)\mathbf{D}\mathbf{C}_k)^{-1}$, as given by

$$\boldsymbol{\Gamma}_k^{-1}(i) = \frac{1}{1-\alpha} \times \left[\boldsymbol{\Gamma}_k^{-1}(i-1) - \frac{\boldsymbol{\Gamma}_k^{-1}(i-1)\boldsymbol{\gamma}_k(i)\boldsymbol{\gamma}_k^{\mathrm{H}}(i)\boldsymbol{\Gamma}_k^{-1}(i-1)}{\frac{1-\alpha}{\alpha} + \boldsymbol{\gamma}_k^{\mathrm{H}}(i)\boldsymbol{\Gamma}_k^{-1}(i)\boldsymbol{\gamma}_k(i)}\right] \quad (38)$$

where $\boldsymbol{\Gamma}_k(i)$ is an estimate of $(\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\hat{\bar{\mathbf{R}}}_k^{-1}(i)\mathbf{D}\mathbf{C}_k)$, and $\boldsymbol{\gamma}_k(i) = \mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{r}_k(i)$, and then, we construct the reduced-rank receiver as

$$\mathbf{w}_k(i) = \hat{\bar{\mathbf{R}}}_k(i)^{-1}\mathbf{D}\mathbf{C}_k\boldsymbol{\Gamma}_k^{-1}(i)\hat{\mathbf{g}}(i). \quad (39)$$

The channel estimate $\hat{\mathbf{g}}(i)$ is obtained through the power method and the RLS technique described in [28]. Following this approach, the SVD on the $L_p \times L_p$ matrix $\mathbf{C}_k^{\mathrm{H}}\mathbf{R}^{-1}(i)\mathbf{C}_k$, as stated in (19) and which requires $O(L_p^3)$, is avoided and replaced by a single matrix–vector multiplication, resulting in the reduction of the corresponding computational complexity on one order of magnitude and no performance loss. In terms of computational complexity, the CMV-RLS algorithm with the interpolated receiver trades off one blind algorithm with complexity $O(M^2)$ against two with complexity $O(M^2/L^2)$ and $O(N_I^2)$ operating in parallel. Since N_I is small as compared to M, it turns out that the new algorithms offer a significant computational advantage over conventional RLS algorithms.

E. Computational Complexity

In this section, we illustrate the computational complexity of the proposed INT structure and algorithms. In Table I, we consider supervised algorithms, whereas the complexity of blind algorithms is depicted in Table II. Specifically, we compare the full-rank, the proposed INT structure, the PD, the PC, and the MWF structures with SG and RLS algorithms.

In general, the INT structure introduces the term M/L, which can reduce the complexity by choosing the decimation factor $L \geqslant 2$. This is relevant for algorithms which have quadratic computational cost with M, i.e., the blind and trained RLS and the blind SG, because the decimation factor L in the denominator favors the proposed scheme which requires complexity $O((M/L)^2)$. This complexity advantage is not verified with linear complexity recursions. For instance, with NLMS algorithms, the proposed INT has a complexity that is slightly superior to the full-rank. Among the other methods, the PD is slightly more complex than the INT. A drawback of PC methods is that they require an SVD with associated cost $O(M^3)$ in order to compute the desired subspace. Although the subspace of interest can be retrieved via computationally efficient tracking algorithms [13], [14], these algorithms are still complex $(O(M)^2)$ and lead to performance degradation as compared to the actual SVD. The MWF technique has a complexity $O(D\bar{M}^2)$, where the variable dimension of the vectors $\bar{M} = M - d$ varies according to the orthogonal decomposition, and the rank $d = 1, \ldots, D$.

In order to illustrate the complexity trend in a comprehensive way, we depict in Fig. 2 curves which describe the computational complexity in terms of the arithmetic operations (additions and multiplications) as a function of the number of parameters M for recursive algorithms. For these curves, we consider $L_p = 6$ and assume that D is equal to $M/2$ for the eigendecomposition approaches. We also include the

TABLE I
COMPUTATIONAL COMPLEXITY OF SUPERVISED ADAPTATION ALGORITHMS

Algorithm	Number of operations per symbol: Additions	Multiplications
LMS-Full-rank	$2M$	$2M+1$
LMS-INT	$2\frac{M}{L}+2N_I$ $+N_IM+(\frac{M}{L})N_I+2$	$3\frac{M}{L}+2N_I$ $+(\frac{M}{L})N_I$
LMS-PC	$O(M^3)+2D$	M^3+2D+1
LMS-PD	$(D-1)^2+2D+1$	D^2+2D+2
MWF-SG	$D(2(\bar{M}-1)^2+\bar{M}+3)$	$D(2\bar{M}^2+5\bar{M}+7)$
RLS-Full-rank	$3(M-1)^2+M^2+2M$	$6M^2+2M+2$
RLS-INT	$3(\frac{M}{L}-1)^2+3(N_I-1)^2$ $+(\frac{M}{L}-1)N_I$ $+N_IM+(\frac{M}{L})^2$ $+N_I^2+2\frac{M}{L}+2N_I$	$6(\frac{M}{L})^2+6N_I^2$ $\frac{M}{L}N_I$ $+3\frac{M}{L}$ $+N_I+2$
RLS-PC	$M^3+3(D-1)^2$ $+D^2+2D$	$O(M^3)+6D^2$ $+2D+2$
RLS-PD	$4(D-1)^2+D^2+2D$	$7D^2+2D+2$
MWF-Recursive	$D(4(\bar{M}-1)^2+2\bar{M})$	$D(4\bar{M}^2+2\bar{M}+3)$

TABLE II
COMPUTATIONAL COMPLEXITY OF BLIND ADAPTATION ALGORITHMS

Algorithm	Number of operations per symbol: Additions	Multiplications
CMV-SG-	M^2+ML_p+2M+1	M^2+ML_p+3M
Full-rank	$(\frac{M}{L})^2+(\frac{M}{L})L_p$	$(\frac{M}{L})^2+(\frac{M}{L})L_p$
CMV-SG-INT	$+N_IM+(\frac{M}{L})N_I$ $+2\frac{M}{L}+N_I+2$	$+\frac{M}{L}N_I+4\frac{M}{L}$ $+N_I$
MWF-SG	$D(2(\bar{M}-1)^2+2)$	$D(2\bar{M}^2+3\bar{M}+5)$
CMV-RLS-Full-rank	$4(M-1)^2+M^2$ $+3(L_p-1)^2-1$ $L_p^2+2L_p+ML_p$	$7M^2+M$ $+L_p^2$ $+ML_p+L_p+4$
CMV-RLS-INT	$4(\frac{M}{L}-1)^2+(\frac{M}{L})^2+L_p^2$ $+3(L_p-1)^2+2\frac{M}{L}L_p$ $+N_IM+3L_p-1$ $+(\frac{M}{L}-1)N_I+(N_I-1)^2$	$7(\frac{M}{L})^2+2\frac{M}{L}$ $+L_p^2+\frac{M}{L}L_p$ $+L_p+2+N_I^2$ $+\frac{M}{L}N_I+N_I$
PC-Wang&Poor	$M^3+2(M-1)^2$	$O(M^3)+2M^2+M$
MWF-Recursive	$D(3(\bar{M}-1)^2+2\bar{M})$	$D(3\bar{M}^2+2\bar{M}+3)$

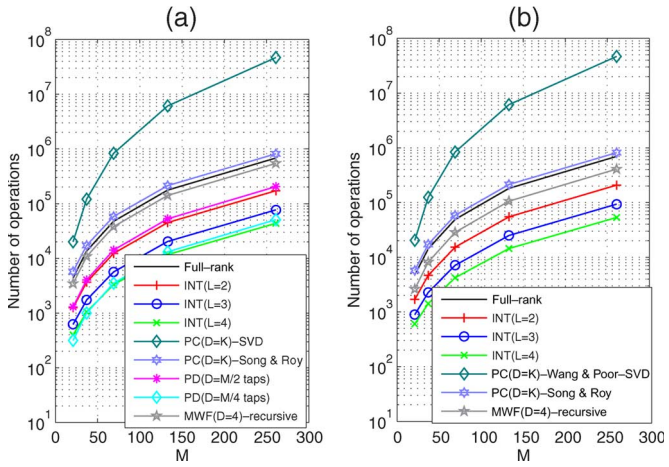

Fig. 2. Complexity in terms of arithmetic operations versus number of received samples (M) for (a) supervised and (b) blind recursive adaptation algorithms.

computational cost of the algorithm of Song and Roy [14], which is capable of significantly reducing the cost required by SVD. In comparison with the existing reduced-rank techniques, the proposed INT scheme is significantly less complex than the PC and the MWF and is slightly less complex than the PD. This is because the analyzed algorithms have quadratic cost (PC with SVD has cubic cost), whereas the INT has complexity $O((M/L)^2)$, as shown in Tables I and II.

V. Global Convergence Properties of the Method and Convergence Analysis of Algorithms

In this section, we discuss the global convergence of the method and its properties, the trajectory of the mean tap vectors, of the excess MSE, and the convergence speed. Specifically, we study the convergence properties of the proposed joint adaptive interpolator and receiver scheme and conclude that it leads to an optimization problem with multiple global minima and no local minima. In this regard and based on the analyzed convergence properties of the method, it suffices to examine the MSE trajectory of only one of the jointly optimized parameter vectors ($\mathbf{w}_k$ or $\mathbf{v}_k$) in order to predict the excess MSE of both blind and supervised adaptive algorithms. We also provide a discussion of the speed of convergence of the INT, as compared to the full-rank.

A. Global Convergence of the Method and Its Properties

1) Interpolated MMSE Design: Let us first consider the trained receiver case and recall the associated MSE expressions in (13) and (14), namely, $J_{\mathrm{MSE}}(\mathbf{v}_k, \boldsymbol{\alpha}(\mathbf{v}_k)) = J(\mathbf{v}_k) = \sigma_{b_k}^2 - \bar{\mathbf{p}}_k^{\mathrm{H}}(i)\bar{\mathbf{R}}_k^{-1}(i)\bar{\mathbf{p}}_k(i)$ and $J_{\mathrm{MSE}}(\boldsymbol{\beta}(\mathbf{w}_k), \mathbf{w}_k) = \sigma_{b_k}^2 - \bar{\mathbf{p}}_{u_k}^{\mathrm{H}}(i)\bar{\mathbf{R}}_{u_k}^{-1}(i)\mathbf{p}_{u_k}(i)$, where $\sigma_{b_k}^2 = E[|b_k(i)|^2]$. Note that points of global minimum of $J_{\mathrm{MSE}}(\mathbf{w}_k(i), \mathbf{v}_k(i)) = E[|b_k(i) - \mathbf{v}_k^{\mathrm{H}}(i)\Re(i)\mathbf{w}_k^*(i)|^2]$ can be obtained by $\mathbf{v}_{\mathrm{opt}} = \arg\min_{\mathbf{v}_k} J(\mathbf{v}_k)$ and $\mathbf{w}_{\mathrm{opt}} = \boldsymbol{\alpha}(\mathbf{v}_{\mathrm{opt}})$ or $\mathbf{w}_{\mathrm{opt}} = \arg\min_{\mathbf{w}_k} J_{\mathrm{MSE}}(\boldsymbol{\beta}(\mathbf{w}_k), \mathbf{w}_k)$ and $\mathbf{v}_{\mathrm{opt}} = \boldsymbol{\beta}(\mathbf{w}_{\mathrm{opt}})$. At a minimum point, $J_{\mathrm{MSE}}(\mathbf{v}_k, \boldsymbol{\alpha}(\mathbf{v}_k))$ equals $J_{\mathrm{MSE}}(\boldsymbol{\beta}(\mathbf{w}_k), \mathbf{w}_k)$, and the MMSE for the proposed structure is achieved. We further note that, since $J(\mathbf{v}_k) = J(t\mathbf{v}_k)$, for every $t \neq 0$, then if $\mathbf{v}_k^\star$ is a point of global minimum of $J(\mathbf{v}_k)$, then $t\mathbf{v}_k^\star$ is also a point of global minimum. Therefore, points of global minimum (optimum interpolator filters) can be obtained by $\mathbf{v}_k^\star = \arg\min_{\|\mathbf{v}_k\|=1} J(\mathbf{v}_k)$. Since the existence of at least one point of global minimum of $J(\mathbf{v}_k)$ for $\|\mathbf{v}_k\| = 1$ is guaranteed by the theorem of Weierstrass [32, Ch. 2, Sec. II-A, App. B], then the existence of (infinite) points of global minimum is also guaranteed for the cost function in (10).

In the context of global convergence, a sufficient but not necessary condition is the convexity, which is verified if its Hessian matrix is positive semidefinite, i.e., $\mathbf{a}^{\mathrm{H}}\mathbf{H}\mathbf{a} \geqslant 0$, for any vector $\mathbf{a}$. First, let us consider the minimization of $J_{\mathrm{MSE}}(\mathbf{w}_k(i), \mathbf{v}_k(i)) = E[|b_k(i) - \mathbf{v}_k^{\mathrm{H}}(i)\Re(i)\mathbf{w}_k^*(i)|^2]$ with fixed interpolators. Such optimization leads to the following Hessian $\mathbf{H} = (\partial/\partial\mathbf{w}_k^{\mathrm{H}})((J_{\mathrm{MSE}}(.))/\partial\mathbf{w}_k) = E[\mathbf{r}_k(i)\mathbf{r}_k^{\mathrm{H}}(i)] = \mathbf{R}_k(i)$, which is positive semidefinite and ensures the convexity of the cost function for the case of fixed interpolators. Let us now consider the joint optimization of the interpolator $\mathbf{v}_k$ and receiver $\mathbf{w}_k$ through an equivalent cost function to (10)

$$\tilde{J}_{\mathrm{MSE}}(\mathbf{z}) = E\left[\left|b - \mathbf{z}_k^{\mathrm{H}}\mathbf{B}\mathbf{z}_k\right|^2\right] \tag{40}$$

where $\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \Re & \mathbf{0} \end{bmatrix}$ is an $(N_I + N/L) \times (N_I + N/L)$ matrix, and the Hessian $(\mathbf{H})$ with respect to $\mathbf{z}_k = [\mathbf{w}_k^{\mathrm{T}}\mathbf{v}_k^{\mathrm{T}}]^{\mathrm{T}}$ is $\mathbf{H} = (\partial/\partial\mathbf{z}_k^{\mathrm{H}})(\partial(\tilde{J}_{\mathrm{MSE}}(.))/\partial\mathbf{z}_k) = E[(\mathbf{z}_k^{\mathrm{H}}\mathbf{B}\mathbf{z}_k - b_k)\mathbf{B}^{\mathrm{H}}] + E[(\mathbf{z}_k^{\mathrm{H}}\mathbf{B}^{\mathrm{H}}\mathbf{z}_k - b_k^*)\mathbf{B}] + E[\mathbf{B}\mathbf{z}_k\mathbf{z}_k^{\mathrm{H}}\mathbf{B}^{\mathrm{H}}] + E[\mathbf{B}^{\mathrm{H}}\mathbf{z}_k\mathbf{z}_k^{\mathrm{H}}\mathbf{B}]$. By examining $\mathbf{H}$, we note that the third and fourth terms yield positive semidefinite matrices ($\mathbf{a}^{\mathrm{H}}E[\mathbf{B}\mathbf{z}_k\mathbf{z}_k^{\mathrm{H}}\mathbf{B}^{\mathrm{H}}]\mathbf{a} \geqslant 0$ and $\mathbf{a}^{\mathrm{H}}E[\mathbf{B}^{\mathrm{H}}\mathbf{z}_k\mathbf{z}_k^{\mathrm{H}}\mathbf{B}]\mathbf{a} \geqslant 0$, $\mathbf{z}_k \neq \mathbf{0}$), whereas the first and second terms are indefinite matrices. Thus, the cost function cannot be classified as convex. However, for a gradient search algorithm, a desirable property of the cost function is that it shows no points of local minimum, i.e., every point of minimum is a point of global minimum (convexity is a sufficient, but not necessary, condition for this property to hold), and it is conjectured that the problem in (40) has this property.

To support this claim, we carried out the following studies.

1) Let us consider the scalar case of the function in (40), which is defined as $f(w, v) = (b - w\,r\,v)^2 = b^2 - 2b\,w\,r\,v + (w\Re\,v)^2$, where r is a constant. By choosing v (the "scalar" interpolator) fixed, it is evident that the resulting function $f(w, v) = (b - w\,c)^2$, where c is a constant is a convex one, whereas for a time-varying interpolator, the curves shown in Fig. 3(a) and (b) indicate that the function is no longer convex but that it also does not exhibit local minima.
2) By taking into account that, for small interpolator filter length N_I $(N_I \leqslant 3)$, $\mathbf{v}_k$ can be expressed in spherical coordinates, and a surface can be constructed. Specifically, we expressed the parameter vector $\mathbf{v}_k$ as follows: $\mathbf{v}_k = r[\cos(\theta)\cos(\phi) \quad \cos(\theta)\sin(\phi) \quad \sin(\theta)]^{\mathrm{T}}$, where r is the radius, θ and ϕ were varied from $-\pi/2$ to $\pi/2$ and $-\pi$ to π, respectively, and (13) was plotted for various scenarios and conditions (SNR, different channels, etc). The plot of the error-performance surface of

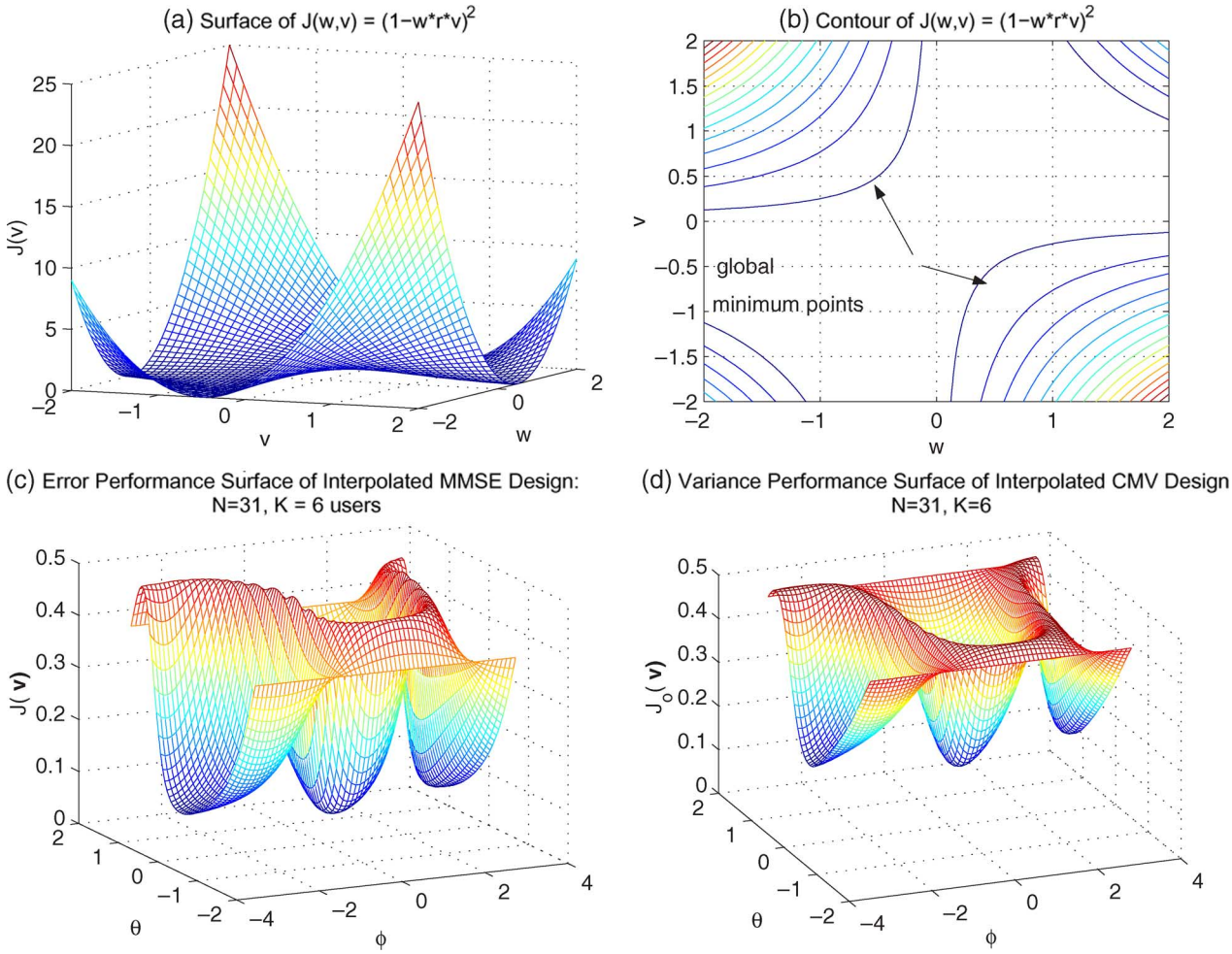

Fig. 3. (a) Error performance surface of the function $f(v,w) = (1 - w * r * v)^2$. (b) Contour plots showing that the function does not exhibit local minima and has multiple global minima. (c) Error performance surface of interpolated MMSE receivers at $E_b/N_0 = 15$ dB for $L = 3$. (d) Variance performance surface of $J_{\rm MV}(\mathbf{v})$ for CMV receivers at $E_b/N_0 = 15$ dB for $L = 2$ and channel with paths given by 0, -6, and -10 dB, spaced by T_c.

$J(\mathbf{v}_k)$, which is depicted in Fig. 3(c), reveals that $J(\mathbf{v}_k)$ has a global minimum value (as it should) but does not exhibit local minima, which implies that (40) has no local minima either. It should be noted that if the cost function in (40) had a point of local minimum, then $J(\mathbf{v}_k)$ in (13) should also exhibit a point of local minimum, even though the reciprocal is not necessarily true: A point of local minimum of $J(\mathbf{v}_k)$ may correspond to a saddle point of $J_{\rm MSE}(\mathbf{v}_k, \mathbf{w}_k)$ if it exists. Note also that the latitude X longitude plot in Fig. 3(c) depicts its two symmetric global minima in the unit sphere.

3) An important feature that advocates the nonexistence of local minima is that the algorithm always converge to the same minimum value, for a given experiment, independently of any interpolator initialization (except for $\mathbf{v}(0) = [0 \cdots 0]^{\rm T}$ that eliminates the signal) for a wide range of SNR values and channels.

2) Interpolated CMV Design: For the blind case, let us first consider the minimization of $J_{\rm MV}(\mathbf{w}_k(i), \mathbf{v}_k(i)) = E[|\mathbf{v}_k^{\rm H}(i)\Re(i)\mathbf{w}_k^*(i)|^2]$ with fixed interpolators subject to $\mathbf{C}_k^{\rm H}\mathbf{D}^{\rm H}\mathbf{w}_k(i) = \mathbf{g}(i)$ and $\|\mathbf{v}_k(i)\| = 1$. It should be noted that global convergence of the CMV method has been established in [6], and in this paper, we treat a similar problem when fixed interpolators are used. Such optimization leads to the following Hessian $\mathbf{H} = (\partial/\partial\mathbf{w}_k^{\rm H})((J_{\rm MV}(.))/\partial\mathbf{w}_k) = E[\mathbf{r}_k(i)\mathbf{r}_k^{\rm H}(i)] = \mathbf{R}_k(i)$, which is positive semidefinite and ensures the convexity of the cost function for the case of fixed interpolators.

Consider the joint optimization of the interpolator $\mathbf{v}_k$ and receiver $\mathbf{w}_k$ via an equivalent cost function to (10)

$$\tilde{J}_{\rm MV}(\mathbf{z}) = E\left[|\mathbf{z}_k^{\rm H}\mathbf{B}\mathbf{z}_k|^2\right] \tag{41}$$

subject to $\mathbf{C}_k^{\rm H}\mathbf{D}^{\rm H}\mathbf{w}_k(i) = \mathbf{g}(i)$, where $\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \Re & \mathbf{0} \end{bmatrix}$ is an $(N_I + N/L) \times (N_I + N/L)$ matrix, and the Hessian ($\mathbf{H}$), with respect to $\mathbf{z}_k = [\mathbf{w}_k^{\rm T}\mathbf{v}_k^{\rm T}]^{\rm T}$, is $\mathbf{H} = (\partial/\partial\mathbf{z}_k^{\rm H})(\partial(\tilde{J}_{\rm MSE}(.))/\partial\mathbf{z}_k) = E[\mathbf{z}_k^{\rm H}\mathbf{B}\mathbf{z}_k\mathbf{B}^{\rm H}] + E[\mathbf{z}_k^{\rm H}\mathbf{B}^{\rm H}\mathbf{z}_k\mathbf{B}] + E[\mathbf{B}\mathbf{z}_k\mathbf{z}_k^{\rm H}\mathbf{B}^{\rm H}] + E[\mathbf{B}^{\rm H}\mathbf{z}_k\mathbf{z}_k^{\rm H}\mathbf{B}]$. By examining $\mathbf{H}$, we note that, as it occurs for the MMSE case, the third and fourth terms yield positive semidefinite matrices ($\mathbf{a}^{\rm H}E[\mathbf{B}\mathbf{z}_k\mathbf{z}_k^{\rm H}\mathbf{B}^{\rm H}]\mathbf{a} \geqslant 0$ and $\mathbf{a}^{\rm H}E[\mathbf{B}^{\rm H}\mathbf{z}_k\mathbf{z}_k^{\rm H}\mathbf{B}]\mathbf{a} \geqslant 0$, $\mathbf{z}_k \neq \mathbf{0}$), whereas the first and second terms are indefinite matrices. Hence, the cost function cannot be classified as convex, although we conjecture that it does not exhibit local minima. Thus, we proceed similarly to the MMSE case to study the surfaces provided by the

problem in (41). Then, we carried out the following studies. 1) We have also plotted the variance performance surface of $J_o(\mathbf{v}_k)$ in (17), depicted in Fig. 3(d). This surface reveals that $J_o(\mathbf{v}_k)$ has a global minimum (as it should) but does not exhibit local minima, which implies that (41) subject to $\mathbf{C}_k^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{w}_k(i) = \mathbf{g}(i)$ has no local minima either. 2) Another important feature that suggests the nonexistence of local minima for the blind algorithms is that they always converge to the same minimum value, for a given experiment, independently of any interpolator initialization (except for $\mathbf{v}(0) = [0 \cdots 0]^{\mathrm{T}}$ that eliminates the signal) for a wide range of parameters.

B. Trajectory of the Mean Tap Vectors

This part is devoted to the analysis of the trajectory of the mean tap vectors of the proposed structure when operating in blind and supervised modes. In our analysis, we employ the so-called independence theory [1], [31, Ch. 9, pp. 390–404] that consists of four points.

1) The received vectors $\mathbf{r}(1), \ldots, \mathbf{r}(i)$ and their interpolated counterparts $\bar{\mathbf{r}}_k(1), \ldots, \bar{\mathbf{r}}_k(i)$ constitute a sequence of statistically independent vectors.
2) At time i, $\mathbf{r}(i)$ and $\bar{\mathbf{r}}_k(i)$ are statistically independent of $b_k(1), \ldots, b_k(i-1)$.
3) At time i, $b_k(i)$ depends on $\mathbf{r}(i)$ and $\mathbf{r}_k(i)$ but is independent of previous $b_k(n)$, for $n = 1, \ldots, i-1$.
4) The vectors $\mathbf{r}(i)$ and $\bar{\mathbf{r}}_k(i)$ and the sample b_k are mutually Gaussian-distributed random variables (r. v.).

In the present context, it is worth noting that the independence assumption holds for synchronous DS-CDMA systems [1], which is the present case, but not for asynchronous models, even though it provides substantial insight.

1) Trained Algorithm: To proceed, let us drop the user k index for ease of presentation and define the tap error vectors $\mathbf{e}_w(i)$ and $\mathbf{e}_v(i)$ at time index i

$$\mathbf{e}_w(i) = \mathbf{w}(i) - \mathbf{w}_{\mathrm{opt}}, \qquad \mathbf{e}_v(i) = \mathbf{v}(i) - \mathbf{v}_{\mathrm{opt}} \tag{42}$$

where $\mathbf{w}_{\mathrm{opt}}$ and $\mathbf{v}_{\mathrm{opt}}$ are the optimum tap vectors that achieve the MMSE for the proposed structure. Substituting the expressions in (42) into (21) and (22), we get

$$\mathbf{e}_w(i+1) = \left[\mathbf{I} - \mu\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\right]\mathbf{e}_w(i) + \mu\bar{\mathbf{r}}(i)e^*(i) \tag{43}$$

$$\mathbf{e}_v(i+1) = \left[\mathbf{I} - \eta\mathbf{u}(i)\mathbf{u}^{\mathrm{H}}(i)\right]\mathbf{e}_v(i) + \eta\mathbf{u}(i)e^*(i). \tag{44}$$

By taking expectations on both sides, we have

$$E\left[\mathbf{e}_w(i+1)\right] = \left[\mathbf{I} - \mu\bar{\mathbf{R}}(i)\right]E\left[\mathbf{e}_w(i)\right] + \mu E\left[\bar{\mathbf{r}}(i)e^*(i)\right] \tag{45}$$

$$E\left[\mathbf{e}_v(i+1)\right] = \left[\mathbf{I} - \eta\mathbf{R}_{\mathbf{u}}(i)\right]E\left[\mathbf{e}_v(i)\right] + \eta E\left[\mathbf{u}(i)e^*(i)\right]. \tag{46}$$

At this point, it should be noted that the two error vectors have to be considered together because of the joint optimization of the interpolator filter and the reduced-rank filter. Rewriting the terms $E[\bar{\mathbf{r}}(i)e^*(i)]$ and $E[\mathbf{u}(i)e^*(i)]$ and using (42) and the independence theory [31, Ch. 9, pp. 390–404], we obtain

$$\begin{aligned} E\left[\bar{\mathbf{r}}(i)e^*(i)\right] = {} & \bar{\mathbf{p}}(i) - E\left[\bar{\mathbf{r}}(i)\mathbf{v}^{\mathrm{T}}(i)\mathfrak{R}^{\mathrm{H}}(i)\right]E\left[\mathbf{e}_w(i)\right] \\ & - E\left[\bar{\mathbf{r}}(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*\right]E\left[\mathbf{e}_v(i)\right] \\ & - E\left[\bar{\mathbf{r}}(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*\mathbf{v}_{\mathrm{opt}}\right] \end{aligned} \tag{47}$$

$$\begin{aligned} E\left[\mathbf{u}(i)e^*(i)\right] = {} & \bar{\mathbf{p}}_{\mathbf{u}}(i) - E\left[\mathbf{u}(i)\mathbf{w}^{\mathrm{T}}(i)\mathfrak{R}^*(i)\right]E\left[\mathbf{e}_v(i)\right] \\ & - E\left[\mathbf{u}(i)\mathbf{v}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^{\mathrm{H}}\right]E\left[\mathbf{e}_w(i)\right] \\ & - E\left[\mathbf{u}(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*\mathbf{v}_{\mathrm{opt}}\right]. \end{aligned} \tag{48}$$

By combining (45)–(48), the trajectory of the error vectors is given by

$$\begin{bmatrix} E\left[\mathbf{e}_w(i+1)\right] \\ E\left[\mathbf{e}_v(i+1)\right] \end{bmatrix} = \mathbf{A}\begin{bmatrix} E\left[\mathbf{e}_w(i)\right] \\ E\left[\mathbf{e}_v(i)\right] \end{bmatrix} + \mathbf{B} \tag{49}$$

where we have the expression shown at the bottom of the next page, and

$$\mathbf{B} = \begin{bmatrix} \mu\bar{\mathbf{p}}(i) - \mu E\left[\bar{\mathbf{r}}(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*\mathbf{v}_{\mathrm{opt}}\right] \\ \eta\bar{\mathbf{p}}_{\mathbf{u}}(i) - \eta E\left[\mathbf{u}(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*\mathbf{v}_{\mathrm{opt}}\right] \end{bmatrix}.$$

Equation (49) implies that the stability of the algorithms in the proposed structure depends on the matrix $\mathbf{A}$. For stability, the convergence factors should be chosen so that the eigenvalues of $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ are less than one.

2) Blind Algorithm: The mean vector analysis of the blind algorithm is slightly different from [6], because our approach uses a decoupled SG channel-estimation technique [28] that yields better channel estimates. Hence, we consider the joint estimation of $\mathbf{w}_k$ and $\mathbf{v}_k$, while $\mathbf{g}$ is a decoupled estimation process. To proceed, let us drop the user k index for ease of presentation and substitute the expressions of (42) into (32) and (34) that gives

$$\begin{aligned} \mathbf{e}_w(i+1) = {} & \left[\mathbf{I} - \mu\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\right]\mathbf{e}_w(i) + \mathbf{DC}(\mathbf{C}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{DC})^{-1}\mathbf{g}(i) \\ & - \mu\mathbf{\Pi}\bar{\mathbf{r}}(i)\mathbf{v}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*(i)\mathbf{w}_{\mathrm{opt}} \\ & - \mu\mathbf{\Pi}\bar{\mathbf{r}}(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^{\mathrm{H}}(i)\mathbf{e}_v(i) \end{aligned} \tag{50}$$

$$\begin{aligned} \mathbf{e}_v(i+1) = {} & \left[\mathbf{I} - \eta\mathbf{u}(i)\mathbf{u}^{\mathrm{H}}(i)\right]\mathbf{e}_v(i) - \eta\mathbf{u}(i)\mathbf{v}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*(i)\mathbf{e}_w(i) \\ & - \eta\mathbf{u}(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*(i)\mathbf{v}_{\mathrm{opt}} \end{aligned} \tag{51}$$

where $\mathbf{\Pi} = \mathbf{I} - \mathbf{DC}(\mathbf{C}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{DC})^{-1}\mathbf{C}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}$, and we used the fact that the scalars have alternative expressions, as $(\mathbf{e}_w^{\mathrm{T}}(i)\mathfrak{R}^{\mathrm{H}}(i)\mathbf{v}_{\mathrm{opt}})^{\mathrm{T}} = (\mathbf{e}_w^{\mathrm{T}}(i)\mathfrak{R}^{\mathrm{H}}(i)\mathbf{v}_{\mathrm{opt}}) = \mathbf{v}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^*(i)\mathbf{e}_w(i)$ and $(\mathbf{e}_v^{\mathrm{T}}(i)\mathfrak{R}^*(i)\mathbf{w}_{\mathrm{opt}})^{\mathrm{T}} = (\mathbf{e}_v^{\mathrm{T}}(i)\mathfrak{R}^*(i)\mathbf{w}_{\mathrm{opt}}) = \mathbf{w}_{\mathrm{opt}}^{\mathrm{T}}\mathfrak{R}^{\mathrm{H}}(i) \times \mathbf{e}_v(i)$. By taking expectations on both sides and eliminating the

term $\mu\mathbf{\Pi}\bar{\mathbf{r}}(i)\mathbf{v}_{\rm opt}\mathfrak{R}^*(i)\mathbf{w}_{\rm opt}$, we get

$$\begin{aligned} E\left[\mathbf{e}_w(i+1)\right] &= \left[\mathbf{I}-\mu\bar{\mathbf{R}}(i)\right]E\left[\mathbf{e}_w(i)\right] \\ &\quad + \mathbf{DC}(\mathbf{C}^{\rm H}\mathbf{D}^{\rm H}\mathbf{DC})^{-1}E\left[\mathbf{g}(i)\right] \\ &\quad - \mu\mathbf{\Pi}E\left[\bar{\mathbf{r}}(i)\mathbf{w}_{\rm opt}^{\rm T}\mathfrak{R}^{\rm H}(i)\right]E\left[\mathbf{e}_v(i)\right] \qquad (52) \\ E\left[\mathbf{e}_v(i+1)\right] &= \left[\mathbf{I}-\eta\mathbf{R}_{\mathbf{u}}(i)\right]E\left[\mathbf{e}_v(i)\right] \\ &\quad - \eta E\left[\mathbf{u}(i)\mathbf{v}_{\rm opt}^{\rm T}\mathfrak{R}^*(i)\right]E\left[\mathbf{e}_w(i)\right] \\ &\quad - \eta E\left[\mathbf{u}(i)\mathbf{w}_{\rm opt}^{\rm T}\mathfrak{R}^*(i)\right]\mathbf{v}_{\rm opt}. \qquad (53) \end{aligned}$$

By combining (52) and (53), the trajectory of the error vectors for the MV case is given by

$$\begin{bmatrix} E\left[\mathbf{e}_w(i+1)\right] \\ E\left[\mathbf{e}_v(i+1)\right] \end{bmatrix} = \mathbf{A}_{\rm MV}\begin{bmatrix} E\left[\mathbf{e}_w(i)\right] \\ E\left[\mathbf{e}_v(i)\right] \end{bmatrix} + \mathbf{B}_{\rm MV} \qquad (54)$$

where

$$\mathbf{A}_{\rm MV} = \begin{bmatrix} \left[\mathbf{I}-\mu\bar{\mathbf{R}}(i)\right] & -\mu\mathbf{\Pi}E\left[\bar{\mathbf{r}}(i)\mathbf{w}_{\rm opt}^{\rm T}\mathfrak{R}^{\rm H}(i)\right] \\ -\eta E\left[\mathbf{u}(i)\mathbf{v}_{\rm opt}^{\rm T}\mathfrak{R}^*(i)\right] & \left[\mathbf{I}-\eta\mathbf{R}_{\mathbf{u}}(i)\right] \end{bmatrix}$$

and

$$\mathbf{B}_{\rm MV} = \begin{bmatrix} \mathbf{DC}(\mathbf{C}^{\rm H}\mathbf{D}^{\rm H}\mathbf{DC})^{-1}E\left[\mathbf{g}(i)\right] \\ -\eta E\left[\mathbf{u}(i)\mathbf{w}_{\rm opt}^{\rm T}\mathfrak{R}^*(i)\right]\mathbf{v}_{\rm opt} \end{bmatrix}.$$

Equation (54) suggests that the stability of the algorithms in the proposed structure depends on the matrix $\mathbf{A}_{\rm MV}$. For stability, the convergence factors should be chosen so that the eigenvalues of $\mathbf{A}_{\rm MV}^{\rm H}\mathbf{A}_{\rm MV}$ are less than one.

C. Trajectory of Excess MSE

Here, we describe the trajectory of the excess MSE at steady state of the trained and the blind SG algorithms.

1) Trained Algorithm: The analysis for the LMS algorithm using the proposed interpolated structure and the computation of its steady-state excess MSE resembles the one in [31, Ch. 9, pp. 390–404]. Here, an interpolated structure with joint optimization of interpolator $\mathbf{v}_k$ and reduced-rank receiver $\mathbf{w}_k$ is taken into account. Despite the joint optimization, for the computation of the excess MSE, one has to consider only the reduced-rank parameter vector $\mathbf{w}_k$ because the MSE attained upon convergence by (13) and (14) should be the same. Here, we will drop the user k index for ease of presentation. Consider the MSE at time $i+1$ as

$$\epsilon(i+1) = E\left[\left|b(i+1)-\mathbf{w}^{\rm H}(i+1)\bar{\mathbf{r}}(i+1)\right|^2\right]. \qquad (55)$$

By using $\mathbf{w}(i+1)=\mathbf{w}_{\rm opt}+\mathbf{e}_w(i+1)$, $\mathbf{w}_{\rm opt}$, $\mathbf{v}_{\rm opt}$ and the fact that the expressions in (13) and (14) are equal for the optimal parameter vectors, the MSE becomes

$$\begin{aligned} \epsilon(i+1) &= \sigma_b^2 - \bar{\mathbf{p}}^{\rm H}(i+1)\bar{\mathbf{R}}^{-1}(i+1)\bar{\mathbf{p}}(i+1) \\ &\quad - \bar{\mathbf{p}}^{\rm H}(i+1)\mathbf{e}_w(i+1) - \mathbf{e}_w^{\rm H}(i+1)\bar{\mathbf{p}}(i+1) \\ &\quad - \mathbf{w}_{\rm opt}^{\rm H}\bar{\mathbf{p}}(i+1) + \mathbf{w}_{\rm opt}^{\rm H}\bar{\mathbf{R}}(i+1)\mathbf{w}_{\rm opt} \\ &\quad + \mathbf{w}_{\rm opt}^{\rm H}\bar{\mathbf{R}}(i+1)\mathbf{e}_w(i+1) \\ &\quad + \mathbf{e}_w^{\rm H}(i+1)\bar{\mathbf{R}}(i+1)\mathbf{w}_{\rm opt} \\ &\quad + E\left[\mathbf{e}_w(i+1)\bar{\mathbf{r}}(i+1)\bar{\mathbf{r}}^{\rm H}(i+1)\mathbf{e}_w^{\rm H}(i+1)\right] \\ &= \sigma_b^2 - \bar{\mathbf{p}}^{\rm H}(i+1)\bar{\mathbf{R}}^{-1}(i+1)\bar{\mathbf{p}}(i+1) \\ &\quad + E\left[\mathbf{e}_w(i+1)\bar{\mathbf{r}}(i+1)\bar{\mathbf{r}}^{\rm H}(i+1)\mathbf{e}_w^{\rm H}(i+1)\right] \\ &= J_{\rm MMSE}(\mathbf{w}_{\rm opt},\mathbf{v}_{\rm opt}) + \xi_{\rm exc}(i+1) \qquad (56) \end{aligned}$$

where $\bar{\mathbf{p}}(i+1)=E[b^*(i+1)\bar{\mathbf{r}}(i+1)]$, $\epsilon_{\min}=J_{\rm MMSE}(\mathbf{w}_{\rm opt},\mathbf{v}_{\rm opt}) = \sigma_b^2-\bar{\mathbf{p}}^{\rm H}(i+1)\bar{\mathbf{R}}^{-1}(i+1)\bar{\mathbf{p}}(i+1)$ is the MMSE achieved by the proposed structure when we have $\mathbf{w}_{\rm opt}$ and $\mathbf{v}_{\rm opt}$, and $\xi_{\rm exc}(i+1)=E[\mathbf{e}_w^{\rm H}(i+1)\bar{\mathbf{r}}(i+1)\bar{\mathbf{r}}^{\rm H}(i+1)\mathbf{e}_w(i+1)]$ is the excess MSE at time $i+1$. To compute the excess MSE, one must evaluate the term $\xi_{\rm exc}(i+1)$. By invoking the independence assumption and the properties of trace [31, Ch. 9, pp. 390–404], we may reduce it as follows:

$$\begin{aligned} &E\left[\mathbf{e}_w^{\rm H}(i+1)\bar{\mathbf{r}}(i+1)\bar{\mathbf{r}}^{\rm H}(i+1)\mathbf{e}_w(i+1)\right] \\ &\qquad = {\rm tr}\left[\bar{\mathbf{R}}(i+1)\mathbf{K}(i+1)\right]. \qquad (57) \end{aligned}$$

In the following steps, we assume that i is sufficiently large such that the matrix $\bar{\mathbf{R}}(i)=\bar{\mathbf{R}}(\infty)=\bar{\mathbf{R}}$. To proceed, let us define some new quantities that will perform a rotation of coordinates to facilitate our analysis, as advocated in [31]. Define $\mathbf{Q}^{\rm H}\bar{\mathbf{R}}\mathbf{Q}=\mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is a diagonal matrix consisting of the eigenvalues of $\bar{\mathbf{R}}$ and $\mathbf{Q}$ is the unitary matrix with the eigenvectors associated with these eigenvalues. Letting $\mathbf{Q}^{\rm H}\mathbf{K}\mathbf{Q}=\mathbf{X}$, we get

$$\begin{aligned} \xi_{\rm exc}(i+1) &= {\rm tr}\left[\bar{\mathbf{R}}\mathbf{K}(i+1)\right] = {\rm tr}\left[\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\rm H}\mathbf{Q}\bar{\mathbf{X}}(i+1)\mathbf{Q}^{\rm H}\right] \\ &= {\rm tr}\left[\mathbf{Q}\mathbf{\Lambda}\bar{\mathbf{X}}(i+1)\mathbf{Q}^{\rm H}\right] = {\rm tr}\left[\mathbf{\Lambda}\bar{\mathbf{X}}(i+1)\right] \qquad (58) \end{aligned}$$

where we used the property of trace, and $\mathbf{Q}^{\rm H}\mathbf{Q}=\mathbf{I}$. Because $\mathbf{\Lambda}$ is a diagonal matrix of dimension M/L, we have

$$\xi_{\rm exc}(i+1) = \sum_{n=1}^{M/L}\lambda_n x_n(i+1) \qquad (59)$$

where x_n, $n=1,2,\ldots,M/L$ are the elements of the diagonal of $\mathbf{X}(i)$. Here, we may use (45) and invoke the independence

$$\mathbf{A} = \begin{bmatrix} (\mathbf{I}-\mu\bar{\mathbf{R}}) - \mu E\left[\bar{\mathbf{r}}(i)\mathbf{v}^{\rm T}(i)\mathfrak{R}^{\rm H}(i)\right] & -\mu E\left[\bar{\mathbf{r}}(i)\mathbf{w}_{\rm opt}^{\rm T}\mathfrak{R}^*(i)\right] \\ -\eta E\left[\mathbf{u}(i)\mathbf{v}_{\rm opt}^{\rm T}\mathfrak{R}^{\rm H}(i)\right] & (\mathbf{I}-\eta\bar{\mathbf{R}}_{\mathbf{u}}) - \eta E\left[\mathbf{u}(i)\mathbf{w}^{\rm T}(i)\mathfrak{R}^*(i)\right] \end{bmatrix}$$

theory [31, Ch. 9, pp. 390–404] in order to describe the correlation matrix of the weight error vector

$$\begin{aligned}\mathbf{K}(i+1) &= E\left[\mathbf{e}_w(i+1)\mathbf{e}_w^{\mathrm{H}}(i+1)\right] \\ &= \left(\mathbf{I}-\mu\bar{\mathbf{R}}(i)\right)\mathbf{K}(i)\left(\mathbf{I}-\mu\bar{\mathbf{R}}(i)\right)+\mu^2\epsilon_{\min}. \quad (60)\end{aligned}$$

Next, using the transformations $\mathbf{Q}^{\mathrm{H}}\bar{\mathbf{R}}\mathbf{Q}=\mathbf{\Lambda}$ and $\mathbf{Q}^{\mathrm{H}}\mathbf{K}\mathbf{Q}=\mathbf{X}$, and similarly to [31, Ch. 9, pp. 390–404], a recursive equation in terms of $\mathbf{X}(i)$ and $\mathbf{\Lambda}$ can be written as

$$\mathbf{X}(i+1)=(\mathbf{I}-\mu\mathbf{\Lambda})\mathbf{X}(i)(\mathbf{I}-\mu\mathbf{\Lambda})+\mu^2\epsilon_{\min}\mathbf{\Lambda} \quad (61)$$

Because of the structure of the above equation, one can decouple the elements $x_n(i)$ from the off-diagonal ones, and thus, $\xi_{\mathrm{exc}}(i+1)$ depends on $x_n(i)$, according to the following recursion:

$$x_n(i+1)=(1-\mu\lambda_n)^2x_n(i)+\mu^2\epsilon_{\min}\lambda_n. \quad (62)$$

At this point, it can be noted that such a recursive relation converges, provided that all the roots lie inside the unit circle, i.e., $(1-\mu\lambda_n)^2<1$ for all n, and thus, we have, for stability

$$0<\mu<\frac{2}{\lambda_{\max}} \quad (63)$$

where $\lambda_{\max}$ is the largest eigenvalue of the matrix $\bar{\mathbf{R}}$. In practice, $\mathrm{tr}[\bar{\mathbf{R}}]$ is used as a conservative estimate of $\lambda_{\max}$. By taking $\lim_{i\to\infty}$ on both sides of (62), we get $x_n(\infty)=(\mu/(2+\mu\lambda_n))\epsilon_{\min}$. Then, taking limits on both sides of (59) and using $x_n(\infty)$, we obtain the expression for the excess MSE at steady state

$$\begin{aligned}\xi_{\mathrm{exc}}(\infty) &= \sum_{n=1}^{M/L}\lambda_n x_n(\infty) \\ &= \sum_{n=1}^{M/L}\frac{\mu\lambda_n}{2+\mu\lambda_n}\epsilon_{\min}=\frac{\frac{\mu}{2}\mathrm{tr}[\bar{\mathbf{R}}]}{1-\frac{\mu}{2}\mathrm{tr}[\bar{\mathbf{R}}]}\epsilon_{\min}. \quad (64)\end{aligned}$$

The expression in (64) can be used to predict semianalytically the excess MSE, where $\bar{\mathbf{R}}$ must be estimated with the aid of computer simulations, since it is a function of the interpolator $\mathbf{v}(i)$. Alternatively, one can conduct the analysis for the interpolator $\mathbf{v}(i)$, which results in the expression $\xi_{\mathrm{exc}}(\infty)=((\eta/2)\mathrm{tr}[\mathbf{R_u}]/(1-(\eta/2)\mathrm{tr}[\mathbf{R_u}]))\epsilon_{\min}$, where η is the step size of the interpolator, the matrix $\mathbf{R}_u=\mathbf{R}_u(\infty)$, and $\mathbf{R}_u(i)=E[\mathbf{u}(i)\mathbf{u}^{\mathrm{H}}(i)]$, as defined in connection with (12). A more complete analytical result, which is expressed as a function of both step sizes μ and η, and statistics of the noninterpolated observation vector $\mathbf{r}(i)$ requires further investigation in order to determine $\mathrm{tr}[\bar{\mathbf{R}}(\infty)]$, which depends on η or $\mathrm{tr}[\mathbf{R_u}(\infty)]$, which depends on μ. Nevertheless, such investigation is beyond the scope of this paper, and it should be remarked that the results would not differ from the semianalytical results derived here [that implicitly take into account the parameters of $\mathbf{v}(i)$].

2) Blind Algorithm: Our algorithm is an MV technique, and its steady-state excess MSE resembles the approach in [6]. In the current context, however, an interpolated structure with joint optimization of interpolator $\mathbf{v}_k$ and reduced-rank receiver $\mathbf{w}_k$ is taken into account. In particular, it suffices to consider, for the computation of the excess MSE, only the reduced-rank parameter vector $\mathbf{w}_k$, because the MSE attained upon convergence by the recursions, that work in parallel for $\mathbf{w}_k$ and $\mathbf{v}_k$, should be the same. Here, we will drop the user k index for ease of presentation. Consider the MSE at time $i+1$ as

$$\epsilon(i+1)=E\left[\left\|b(i+1)-\mathbf{w}^{\mathrm{H}}(i+1)\bar{\mathbf{r}}(i+1)\right\|^2\right]. \quad (65)$$

By using $\mathbf{w}(i+1)=\mathbf{w}_{\mathrm{opt}}+\mathbf{e}_w(i+1)$ and the independence assumption, the MSE becomes

$$\begin{aligned}\epsilon(i+1) =& \epsilon_{\min}-E\left[b(i+1)\bar{\mathbf{r}}^{\mathrm{H}}(i+1)\right]\mathbf{e}_w(i+1) \\ &-\mathbf{e}_w^{\mathrm{H}}(i+1)E\left[b^*(i+1)\bar{\mathbf{r}}(i+1)\right] \\ &+\mathbf{w}_{\mathrm{opt}}^{\mathrm{H}}\bar{\mathbf{R}}(i+1)\mathbf{e}_w(i+1) \\ &+\mathbf{e}_w^{\mathrm{H}}(i+1)\bar{\mathbf{R}}(i+1)\mathbf{w}_{\mathrm{opt}}+\xi_{\mathrm{exc}}(i+1) \quad (66)\end{aligned}$$

where $\epsilon_{\min}=\sigma_b-E[b(i+1)\bar{\mathbf{r}}^{\mathrm{H}}(i+1)]\mathbf{w}_{\mathrm{opt}}-\mathbf{w}_{\mathrm{opt}}^{\mathrm{H}}E[b^*(i+1)\bar{\mathbf{r}}(i+1)]+\mathbf{w}_{\mathrm{opt}}^{\mathrm{H}}\bar{\mathbf{R}}(i+1)\mathbf{w}_{\mathrm{opt}}$ is the MSE with the optimal reduced-rank receiver $\mathbf{w}_{\mathrm{opt}}$, and the optimal interpolator $\mathbf{v}_{\mathrm{opt}}$ and $\xi_{\mathrm{exc}}(i+1)=E[\mathbf{e}_w^{\mathrm{H}}(i+1)\bar{\mathbf{R}}(i+1)\mathbf{e}_w(i+1)]$ is the excess MSE at time $i+1$. Since $\lim_{i\to\infty}E[\mathbf{e}_w(i)]=\mathbf{0}$, we have

$$\lim_{i\to\infty}\epsilon(i+1)=\epsilon_{\min}+\lim_{i\to\infty}\xi_{\mathrm{exc}}(i+1). \quad (67)$$

Note that the second term in (67) is the steady-state excess MSE due to adaptation, which is denoted by $\bar{\xi}_{\mathrm{exc}}$ and which is related to $\mathbf{w}$ by

$$\bar{\xi}_{\mathrm{exc}}(\infty)=\lim_{i\to\infty}\mathrm{tr}E\left[\bar{\mathbf{R}}\mathbf{e}_w(i+1)\mathbf{e}_w^{\mathrm{H}}(i+1)\right]. \quad (68)$$

Let us define $\mathbf{R}_e(i)=E[\mathbf{e}_w(i)\mathbf{e}_w^{\mathrm{H}}(i)]$ and $\mathbf{R}_e=\lim_{i\to\infty}\mathbf{R}_e(i)$, and use the property of trace to obtain

$$\bar{\xi}_{\mathrm{exc}}(\infty)=\mathrm{tr}E[\bar{\mathbf{R}}\mathbf{R}_e]=\mathrm{vec}^{\mathrm{H}}(\bar{\mathbf{R}})\mathrm{vec}(\mathbf{R}_e). \quad (69)$$

At this point, it can be noted that to assess $\bar{\xi}_{\mathrm{exc}}(\infty)$, it is sufficient to study $\mathbf{R}_e$, which depends on the trajectory of the tap error vector. For simplicity and similarly to [6], we assume that $\mathbf{e}_g(i)\approx\mathbf{C}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{e}_w(i)$, which is valid as the adaptation approaches steady state. Using the expression of $\mathbf{e}_w(i+1)$ and taking expectation on both sides of $\mathbf{e}_w(i+1)\mathbf{e}_w^{\mathrm{H}}(i+1)$, the resulting matrix $\mathbf{R}_e(i+1)$ becomes

$$\begin{aligned}&\mathbf{R}_e(i+1) \\ &\approx\mathbf{R}_e(i)-\mu\left(\mathbf{R}_e(i)\bar{\mathbf{R}}(i)\mathbf{\Pi}+\mathbf{\Pi}\bar{\mathbf{R}}\mathbf{R}_e(i)\right) \\ &\quad-\mu E\left[\mathbf{\Pi}\mathbf{e}_w(i)\mathbf{w}_{\mathrm{opt}}^{\mathrm{H}}\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\mathbf{\Pi}+\mathbf{\Pi}\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\mathbf{w}_{\mathrm{opt}}\mathbf{e}_w^{\mathrm{H}}(i)\mathbf{\Pi}\right] \\ &\quad+\mu^2E\left[\mathbf{\Pi}\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)(\mathbf{w}_{\mathrm{opt}}\mathbf{w}_{\mathrm{opt}}^{\mathrm{H}}+\mathbf{R}_e(i))\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\mathbf{\Pi}\right] \quad (70)\end{aligned}$$

where $\mathbf{\Pi}=\mathbf{I}-\mathbf{D}\mathbf{C}(\mathbf{C}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\mathbf{D}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}$. Since $\lim_{i\to\infty}\mathbf{R}_e(i+1)=\mathbf{R}_e$ and $\lim_{i\to\infty}E[\mathbf{e}_w(i)=\mathbf{0}$, taking limits on both sides of (70) yields

$$\begin{aligned}&\mathbf{R}_e\bar{\mathbf{R}}\mathbf{\Pi}+\mathbf{\Pi}\bar{\mathbf{R}}\mathbf{R}_e \\ &\approx\mu E\left[\mathbf{\Pi}\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\left(\mathbf{w}_{\mathrm{opt}}\mathbf{w}_{\mathrm{opt}}^{\mathrm{H}}+\mathbf{R}_e(i)\right)\bar{\mathbf{r}}(i)\bar{\mathbf{r}}^{\mathrm{H}}(i)\mathbf{\Pi}\right]. \quad (71)\end{aligned}$$

Here, an expression for $\bar{\xi}_{\rm exc}(\infty)$ can be obtained by using the properties of the Kronecker product and arranging all elements of a matrix into a vector columnwise through the vec operation. Hence, the expression for the steady-state excess MSE becomes

$$\bar{\xi}_{\rm exc}(\infty) = {\rm tr}[\bar{\mathbf{R}}\mathbf{R}_e] = \mu {\rm vec}^{\rm H}(\bar{\mathbf{R}})\mathbf{T}^{-1}\mathbf{a} \tag{72}$$

where $\mathbf{T}=(\bar{\mathbf{R}}\mathbf{\Pi})^{\rm T}\otimes\mathbf{I}+\mathbf{I}\otimes(\mathbf{\Pi}\bar{\mathbf{R}})-\mu[\mathbf{\Pi}^{\rm T}\otimes\mathbf{\Pi}]E[(\mathbf{r}(i)\mathbf{r}(i)^{\rm H})^{\rm T}\otimes(\mathbf{r}(i)\mathbf{r}(i)^{\rm H})]$, $\mathbf{a} = [(\mathbf{\Pi})^{\rm T}\otimes\mathbf{\Pi}]\ E[(\mathbf{r}(i)\mathbf{r}^{\rm H}(i))^{\rm T}\otimes(\mathbf{r}(i)\mathbf{r}^{\rm H}(i))]$ ${\rm vec}(\mathbf{w}_{\rm opt}\mathbf{w}_{\rm opt}^{\rm H})$, and $\otimes$ accounts for the Kronecker product. The expression in (72) can be used to predict, semianalytically, the excess MSE, where the matrices $\bar{\mathbf{R}}$ and $\mathbf{T}$ and the vector $\mathbf{a}$ are computed through simulations.

D. Transient Analysis and Convergence Speed

With regard to convergence speed, adaptive receivers/filters have a performance which is proportional to the number of adaptive elements M [1], [19], [31]. Assuming stationary noise and interference, full-rank schemes with RLS algorithms take $2M$ iterations to converge, while SG algorithms require at least an order of magnitude more iterations than RLS techniques [31]. In addition, it is expected that RLS methods do not show excess MSE (when $\alpha = 1$ and operating in a stationary environment), and its convergence is independent of the eigenvalues of the input correlation matrix.

With the proposed INT reduced-rank scheme, the convergence can be made faster due to the reduced number of filter coefficients, and the decimation factor L can be varied in order to control the speed and ability of the filter to track changing environments. As given in the Appendix, we mathematically explain how the INT structure can obtain gains in convergence speed over full-rank schemes with SG and RLS algorithms, respectively.

For SG algorithms, the analysis of the transient components given in the Appendix of the INT scheme reveals that the speed of convergence depends on the eigenvalue spread of the reduced-rank covariance matrix. In principle, we cannot mathematically guarantee that the INT always converges faster than the full-rank, but several studies that examine the eigenvalue spread of the full-rank and the INT covariance matrix show that, for the same data, the INT structure is able to consistently reduce the eigenvalue spread found in the original data covariance matrix, thus explaining its faster convergence in all analyzed scenarios.

For RLS techniques, the analysis of the transient components given in the Appendix guarantees mathematically that the INT is able to converge faster due to the reduced number of filter elements, and we show that the INT with the RLS converges in about $2M/L$ iterations, as compared to the full-rank, which requires $2M$ iterations.

VI. Simulations

In this section, we investigate the effectiveness of the proposed linear-receiver structure and algorithms via simulations and verify the validity of the convergence analysis undertaken for predicting the MSE obtained by the adaptive algorithms. We have conducted experiments under stationary and nonstationary scenarios to assess the convergence performance in terms of signal-to-interference-plus-noise ratio (SINR) of the proposed structure and algorithms and compared them with other recently reported techniques, namely, adaptive versions of the MMSE [19] and CMV [6] full-rank methods, the eigendecomposition (PC) [12], [13], the PD [15], and the MWF [18] reduced-rank techniques with rank D. Moreover, the bit-error-rate (BER) performance of the receivers employing the analyzed techniques is assessed for different loads, processing gains (N), channel paths (L_p) and profiles, and fading rates. The DS-CDMA system employs Gold sequences of length $N = 31$ and $N = 63$.

Because we focus on the downlink, users experiment under the same channel conditions. All channels assume that $L_p = 6$ as an upper bound (even though the effective number of paths will be indicated in the experiments). For fading channels, the channel coefficients $h_l(i) = p_l\alpha_l(i)(l = 0, 1, 2)$, where $\sum_{l=1}^{L_p} p_l^2 = 1$, and $\alpha_l(i)$ is a complex unit variance Gaussian random sequence obtained by passing complex white Gaussian noise through a filter with approximate transfer function $c/\sqrt{1-(f/f_d)^2}$, where c is a normalization constant, $f_d = v/\lambda_c$ is the maximum Doppler shift, λ_c is the wavelength of the carrier frequency, and v is the speed of the mobile [33]. This procedure corresponds to the generation of independent sequences of correlated unit power Rayleigh r. v. $(E[|\alpha_l^2(i)|] = 1)$. The phase ambiguity derived from the blind channel-estimation method in [28] is eliminated in our simulations by using the phase of $\mathbf{g}(0)$ as a reference to remove the ambiguity, and for fading channels, we assume ideal phase tracking and express the results in terms of the normalized Doppler frequency f_dT (cycles per symbol). Alternatively, differential modulation can be used to account for the phase rotation. For the proposed interpolated receivers structures, we employ $M = (N + L_p - 1)/L$ adaptive elements for $L = 2, 3, 4,$ and 8, and when M is not an integer, we will approximate it to the nearest integer. For the full-rank receiver, we have $M = (N + L_p - 1)$.

In the following experiments, the type of adaptive algorithms used and their mode of operation, i.e., training mode, decision-directed mode, and blind mode, are indicated. For the training-based algorithms, the receiver employs training sequences with N_{tr} symbols and then switches to decision-directed mode. The full-rank receiver is considered with the NLMS and RLS techniques, the interpolated receivers are denoted INT, the PC method [12] requires an SVD on the full-rank covariance matrix, and the subspace dimension is chosen as $D = K$. For the PD approach, the columns of the projection matrix are nonoverlapping segments of $\mathbf{s}_k$, as described in [15], whereas, for the MWF and its SG and recursive adaptive versions (MWF-SG and MWF-rec) [18], the number of stages D is optimized for each scenario. The RAKE receiver in supervised mode uses the NLMS and the RLS techniques and the training sequence in order to estimate its parameters. With respect to blind algorithms and the full-rank receiver, the SG algorithm corresponds to the one in [6] with a normalized step size similar to the one introduced in Section IV-A, and the RLS corresponds to the one reported in [6]. The proposed interpolated receiver, i.e., the

INT, uses the CMV-SG and CMV-RLS algorithms particularly designed for it. The different receiver techniques, algorithms, processing gain N, the decimation factor L, and other parameters are depicted in the legends. The eigendecomposition-based receiver of Wang and Poor [13] is denoted subspace W and P and employs an SVD to compute its eigenvectors and eigenvalues. With regard to blind channel estimation, we employ the method in [28] for all SG-based receivers, whereas for the RLS-based receivers, we adopt the study in [28]. The blind MWF and its adaptive versions (blind MWF-SG and blind MWF-rec) [18] have their rank D optimized for each situation and employ the blind channel estimation in [28] to obtain the effective signature sequence in multipath. For the RAKE receiver [33], we also employ the SG blind channel estimation of [28] when compared to other SG-based multiuser receivers, whereas for the comparison with RLS-based receivers, we use its RLS version [28].

A. MSE Convergence Performance: Analytical Results

Here, we verify that the results (64) and (72) of the section on convergence analysis of the mechanisms can provide a means of estimating the excess MSE. The steady-state MSE between the desired and the estimated symbol obtained through simulation is compared with the steady-state MSE computed via the expressions derived in Section VI. In order to illustrate the usefulness of our analysis, we have carried out some experiments. The interpolator filters were designed with $N_I = 3$ elements, and the channels have three paths with gains 0, -6, and -10 dB, respectively, where, in each run, the delay of the second path (τ_2) is given by a discrete uniform r. v. between one and four chips, and the third path is computed with a discrete uniform r. v. between one and $(5 - \tau_2)$ chips in a scenario with perfect power control.

In the first experiment, we have considered the LMS algorithm in trained mode and tuned the parameters of the mechanisms, in order to achieve a low steady-state MSE upon convergence. The parameters of convergence, i.e., μ, are 0.05, 0.06, 0.075, and 0.09 for the full-rank and the INT with L $= 2, 3,$ and 4, respectively, and $\eta = 0.005$ for the interpolator with all L. The results are shown in Fig. 4(a) and indicate that the analytical results closely match those obtained through simulation upon convergence, verifying the validity of our analysis.

In the second experiment, we have considered the blind SG algorithm and tuned the parameters of the mechanisms, in order to achieve a low steady-state MSE upon convergence, similarly to the LMS case. The chosen values for μ are 0.0009, 0.001, 0.0025, and 0.004 for the full-rank and the INT with L $= 2, 3,$ and 4, respectively, and $\eta = 0.005$ for the interpolator with all L. The curves depicted in Fig. 4(b) reveal that a discrepancy is verified in the beginning of the convergence process, when the estimated covariance matrix is constructed with few samples. In addition, this mismatch between the theoretical and simulated curves is explained by the fact that blind algorithms are noisier than trained techniques [5]. However, as time goes by and the data record is augmented, the statistics of the signals is acquired, and the modeled and simulated MSE curves come to a greater agreement.

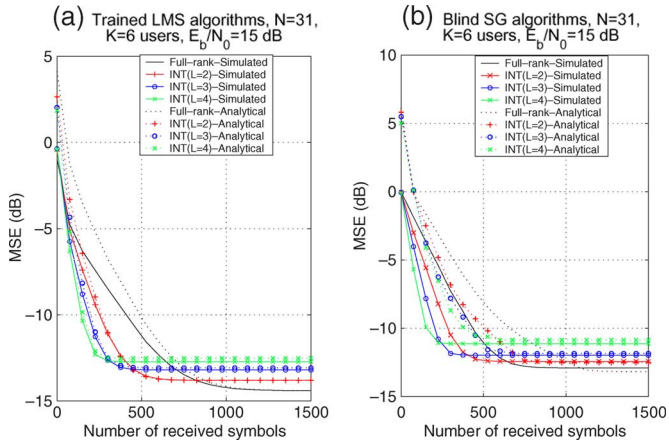

Fig. 4. MSE convergence for analytical and simulated results versus number of received symbols using (a) trained LMS algorithms and (b) blind SG algorithms.

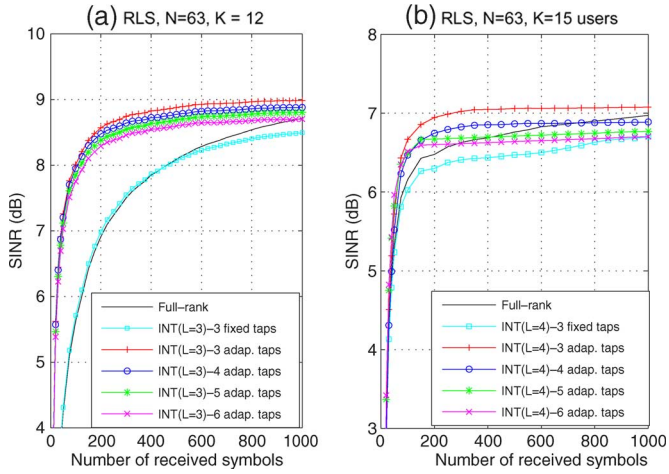

Fig. 5. Design of interpolator filters to obtain the best dimensions for N_I with random three-path channel parameters (r. v. between -1 and 1) as given in Section IV-A, where the scenario has equal power users. (a) Trained RLS-type algorithms at $E_b/N_0 = 12$ dB. (b) Blind CMV-RLS-type algorithms at $E_b/N_0 = 15$ dB.

B. SINR Convergence Performance

The SINR at the receiver end is used here to assess the convergence performance of the analyzed methods. In the following experiments, we will assess the SINR performance of the analyzed adaptive receiver techniques and their corresponding algorithms, namely, the proposed interpolated receiver, the PC, the PD, the MWF, and the RAKE. We remark that the parameters of the algorithms have been tuned in order to optimize performance, and the receiver parameters have been carefully chosen to provide a fair comparison among the analyzed methods.

First, let us consider the issue of how long should be the interpolator filter. Indeed, the design of the interpolator filter is a fundamental issue in our approach because it affects its convergence and BER performance. In order to obtain the most adequate dimension for the interpolator filter $\mathbf{v}_k$, we conducted experiments with values ranging from $N_I = 3$ to $N_I = 6$, which correspond to the ones shown in Fig. 5 for the supervised and blind modes with the RLS, respectively. The

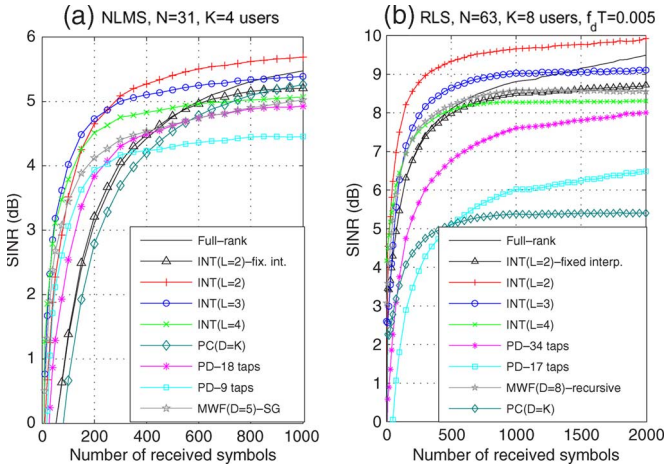

Fig. 6. SINR performance of the receivers with (a) NLMS, $E_b/N_0 = 8$ dB, and three interferers with power levels 7 dB above the desired user and channel parameters $p_0 = 1$, $p_2 = 0.5$, and $p_4 = 0.3$ (spaced by $2T_c$). (b) RLS, $E_b/N_0 = 12$ dB, and three interferers with power levels 10 dB above the desired user with fading and channel parameters $p_0 = 1$, $p_2 = 0.7$, and $p_4 = 0.5$ (spaced by $2T_c$).

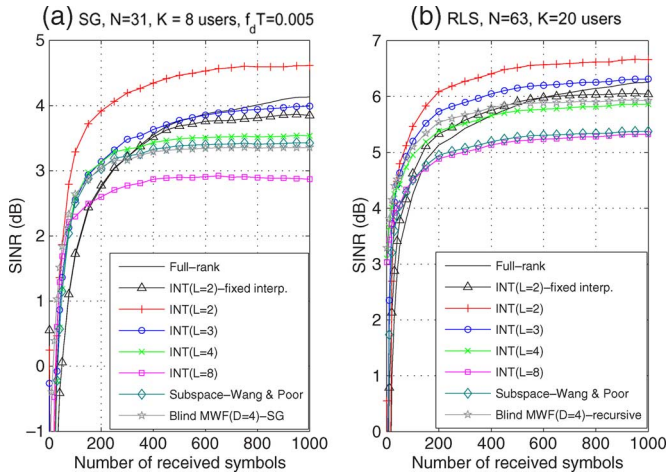

Fig. 7. SINR performance of (a) blind SG algorithms with channel parameters $p_0 = 1$, $p_2 = 0.5$, and $p_4 = 0.5$ (spaced by $2T_c$) where two interferers work at a power level 7 dB above the desired user that operates at $E_b/N_0 = 15$ dB with (b) fading RLS algorithm with $E_b/N_0 = 15$ dB without fading, the three-path channel parameters are random, as in Section IV-A, and the received powers of the interferers are log-normal r. v. with associated standard deviation 3 dB.

results indicate that SINR performance was not sensitive to an increase in the number of taps in $\mathbf{v}_k$, and the best results for all algorithms were obtained with $N_I = 3$. For this reason and to keep the complexity low, we selected $N_I = 3$ for the remaining experiments. We also remark that the simulation-aided design of the interpolator dimension was carried out for systems with different N, K, L, channel profiles, and fading rates, indicating that $N_I = 3$ is a satisfactory dimension. The SINR convergence curves show that the proposed structure with adaptive interpolators is considerably superior to the fixed interpolator approach and to the full-rank receiver.

Fig. 6 illustrates experiments where the INT is compared to other reduced-rank techniques in training and decision-directed modes. In both experiments, a training sequence is provided to the receivers with 200 symbols, and then, the algorithms switch to decision-directed mode. The parameters of the receivers for all methods were optimized, and the results show that the proposed structure with adaptive interpolators and $L = 2$ achieves the best performance and is significantly superior to the INT with a fixed interpolator. The convergence performance of the INT for various L is superior to the full-rank one and to the PC and PD methods. The PC method performs well when K is small, but it is outperformed, both in terms of convergence speed and final SINR, by the INT with $L = 2$ and 3. The INT with $L = 3$ and $L = 4$ are also superior to the PD method with 18 and nine elements, whereas the INT with $L = 4$ has a performance comparable with the MWF adaptive versions.

In Fig. 7, the SINR performance of the analyzed receivers is examined in blind mode. The parameters of the receivers for all methods were optimized, and the results show that the proposed structure with adaptive interpolators and $L = 2$ achieves the best performance. The convergence performance of the novel structure for various L is superior to the full-rank one and to the other methods. Note that subspace approach of Wang and Poor performs very well for small K, but when K is larger, its performance degrades considerably. The INT shows very good

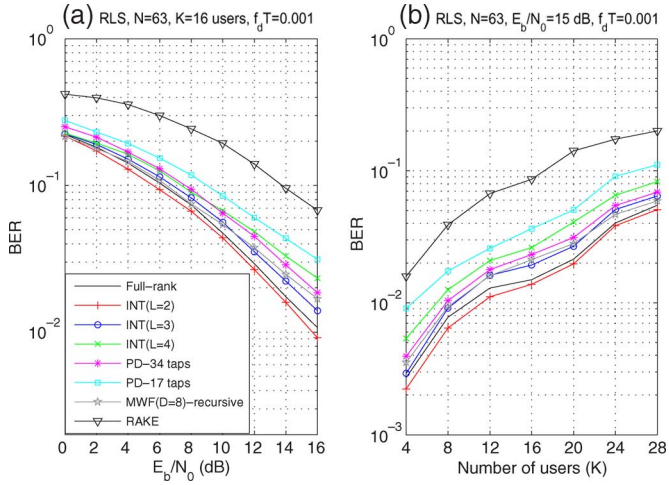

Fig. 8. BER performance of trained RLS algorithms versus (a) E_b/N_0 and (b) number of users.

performance in all situations and requires lower computational costs than the other techniques.

C. BER Performance

In this section, the BER performance of the different receiver techniques is investigated. In Fig. 8, the BER curves for the RLS algorithms in trained and decision-directed modes are shown. The channel parameters are $p_0 = 1$, $p_1 = 0.7$, and $p_2 = 0.5$, where, in each run, the delay of the second path (τ_2) is given by a discrete uniform r. v. between one and four chips, and the third path is computed with a discrete uniform r. v. between 1 and $(5 - \tau_2)$ chips. In these experiments, the received powers of the interferers are log-normal r. v. with associated standard deviation 3 dB. We remark that the proposed methods also perform well with other channel profiles and fading rates. The receivers are trained with 200 symbols and are then switched to decision-directed mode and process 2000 symbols, averaged

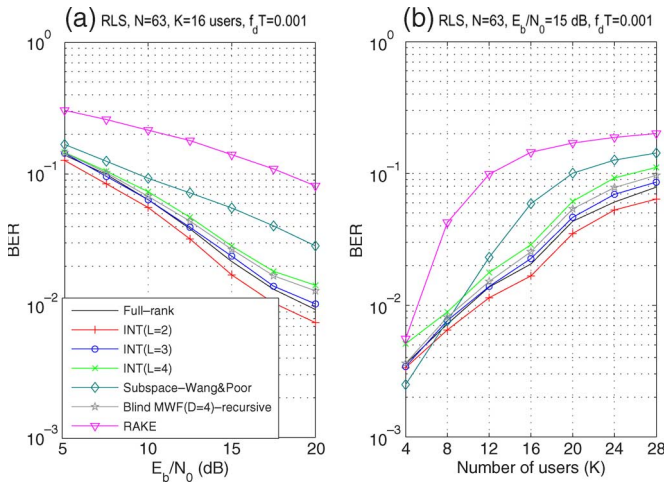

Fig. 9. BER performance of blind RLS algorithms versus (a) E_b/N_0 and (b) number of users.

over 200 experiments with optimized parameters for each scenario. The results show that the INT with $L = 2$ achieves the best performance, followed by the full-rank receiver, the INT with $L = 3$, the MWF, the PD approach, the INT with $L = 4$, the PC, and the RAKE receiver.

In Fig. 9, the BER curves for the RLS-type algorithms in blind mode, respectively, are shown. The receivers process 2000 symbols, averaged over 200 experiments with optimized parameters for each scenario. In these simulations, the received powers of the interferers are log-normal r. v. with associated standard deviation of 3 dB. The results show that the INT with $L = 2$ achieves the best performance, followed by the full-rank receiver, the INT with $L = 3$, the MWF, the INT with $L = 4$, the subspace receiver of Wang and Poor, and the RAKE receiver. Note that the receivers can accommodate more users and cope with larger systems when working with RLS-type algorithms and that the INT structure with $L = 4$ outperforms the RAKE and Wang and Poor's (for $K \geqslant 8$) receivers, the INT with $L = 2$ outperforms the full-rank receiver, and the INT with $L = 3$ has a very close performance to the full-rank. The blind MWF versions are slightly inferior to the INT with $L = 3$ and suffer from the fact that tridiagonalization does not occur, deteriorating its performance.

VII. Conclusion

We proposed adaptive reduce-rank receivers for DS-CDMA based on interpolated FIR filters with adaptive interpolators. The novel receiver structure and algorithms were assessed in various scenarios, outperforming previously reported techniques with a very attractive tradeoff between performance and complexity. An analysis of the convergence properties of the method was undertaken, indicating that the novel cost function does not exhibit local minima. Furthermore, a convergence analysis of the algorithms was shown to be valid in predicting the excess MSE upon convergence for the blind and trained SG algorithms. In terms of computational complexity, the AIFIR receivers are simpler than the full-rank receiver, much simpler than reduced-rank eigendecomposition techniques, and compete favorably with the MWF. The BER performance of the interpolated receivers is superior to the subspace receiver and the MWF and is close to the full-rank one, even with $L = 4$. Finally, with respect to convergence, the proposed receivers exhibit a faster response and greater flexibility than other analyzed methods, since the designer can choose the decimation factor, depending on the need for faster convergence or higher steady-state performance.

Appendix A
Proof of Lemma in Section IV-D

Let $\mathbf{R}$ be a positive semidefinite Hermitian symmetric matrix and its eigenvalues be ordered as $\lambda_{\max} = \lambda_1 > \lambda_2 \geqslant \cdots \geqslant \lambda_{N-1} > \lambda_N = \lambda_{\min} \geqslant 0$ with corresponding eigenvectors $\mathbf{q}_m$ $(m = 1, 2, \ldots, N)$. Consider an initial vector $\hat{\mathbf{v}}(0) = \sum_{m=1}^{N} c_m \mathbf{q}_m$, where c_m are scalars with $c_1 = \hat{\mathbf{v}}^{\mathrm{H}}(0)\mathbf{q}_1 \neq 0$. Using the power iterations, we have

$$\begin{aligned}\hat{\mathbf{v}}(i) &= \mathbf{R}\hat{\mathbf{v}}(i-1) = \mathbf{R}^i\hat{\mathbf{v}}(0)\\ &= c_1\lambda_1^i\mathbf{q}_1 + c_2\lambda_2^i\mathbf{q}_2 + \cdots + c_N\lambda_N^i\mathbf{q}_N\\ &= c_1\lambda_1^i(\mathbf{q}_1 + c_2/c_1(\lambda_2/\lambda_1)^k\mathbf{q}_2\\ &\quad + \cdots + c_N/c_1(\lambda_N/\lambda_1)^i\mathbf{q}_N. \end{aligned} \tag{73}$$

If we normalize the above equation, we obtain

$$\frac{\hat{\mathbf{v}}(i)}{\|\hat{\mathbf{v}}(i)\|} = \frac{(\mathbf{q}_1 + c_2/c_1(\lambda_2/\lambda_1)^k\mathbf{q}_2 + \cdots + c_N/c_1(\lambda_N/\lambda_1)^i\mathbf{q}_N}{\|(\mathbf{q}_1 + c_2/c_1(\lambda_2/\lambda_1)^k\mathbf{q}_2 + \cdots + c_N/c_1(\lambda_N/\lambda_1)^i\mathbf{q}_N\|} \tag{74}$$

and, since $0 \leqslant (\lambda_m/\lambda_1) < 1$, for $2 \leqslant m \leqslant N$, then $\lim_{i\to\infty}(\lambda_m/\lambda_1)^i = 0$ for $2 \leqslant m \leqslant N$. Thus, we conclude that

$$\lim_{i\to\infty} \frac{\hat{\mathbf{v}}(i)}{\|\hat{\mathbf{v}}(i)\|} = \mathbf{q}_{\max} = \mathbf{q}_1. \tag{75}$$

Now, let us make $\mathbf{A} = \mathbf{I} - \nu\mathbf{R}$, where $\nu = 1/\mathrm{tr}[\mathbf{R}]$ and whose eigenvalues are $\lambda'_m = 1 - (\lambda_m/\mathrm{tr}[\mathbf{R}])$, $m = 1, \ldots, N$. Since $\mathrm{tr}[\mathbf{R}] = \sum_{m=1}^{M}\lambda_m \geqslant \lambda_1 = \lambda_{\max}$, then $0 \leqslant \lambda'_1 \leqslant \lambda'_2 \leqslant \cdots \leqslant \lambda'_{N-1} < \lambda'_N$. Therefore, from the development in (73)–(75), we have that the recursion $\hat{\mathbf{v}}(i) = (\mathbf{I} - \nu\hat{\mathbf{R}})\hat{\mathbf{v}}(i-1)$, $i = 1, 2, \ldots$ results in

$$\lim_{i\to\infty} \frac{\hat{\mathbf{v}}(i)}{\|\hat{\mathbf{v}}(i)\|} = \mathbf{q}'_N \tag{76}$$

where $\mathbf{q}'_N$ is the normalized eigenvector of $\mathbf{A}$ associated with $\lambda'_{\max} = \lambda'_N = 1 - (\lambda_{\min}/\mathrm{tr}[\mathbf{R}_{\mathbf{u}_k}])$, i.e., $(\mathbf{I} - \nu\hat{\mathbf{R}})\mathbf{q}'_N = (1 - \nu\lambda_{\min})\mathbf{q}'_N$, and hence, $\mathbf{q}'_N = \mathbf{q}_{\min}$.

Appendix B
Convergence Speed of the INT Scheme With SG Algorithms

In this Appendix, we assess the convergence speed of the proposed INT receiver scheme through the transient component analysis of SG algorithms. By using a similar analysis to the study in [19] and [31, Ch. 9, pp. 390–404; App. J,

pp. 924–927] (with the solution to differential equations), let us express the excess MSE in (59) as a function of its transient and steady-state components

$$\begin{aligned}\xi_{\rm exc}(i) &= \sum_{n=1}^{M/L} \lambda_n x_n(i) = \boldsymbol{\lambda}^{\rm H}\mathbf{x}(i) \\ &= \sum_{n=1}^{M/L} \bar{c}_n^{\,i}\boldsymbol{\lambda}^{\rm H}\mathbf{g}_n\mathbf{g}_n^{\rm H}\left[\mathbf{x}(0)-\mathbf{x}(\infty)\right] + \xi_{\rm exc}(\infty) \\ &= \xi_{\rm trans}(i) + \xi_{\rm exc}(\infty) \end{aligned} \tag{77}$$

where $\bar{c}_n$ is the nth eigenvalue of an $M/L \times M/L$ matrix $\mathbf{T}$ whose entries are

$$t_{nj} = \begin{cases} (1-\mu\lambda_n)^2 & n = j \\ \mu^2\lambda_n\lambda_j & n \neq j \end{cases}. \tag{78}$$

According to the above equation, the speed of convergence of the proposed INT structure for SG algorithms is given by the transient component $\xi_{\rm trans}(i) = \sum_{n=1}^{M/L} \bar{c}_n^{\,i}\boldsymbol{\lambda}^{\rm H}\mathbf{g}_n\mathbf{g}_n^{\rm H}[\mathbf{x}(0) - \mathbf{x}(\infty)]$, which can be alternatively expressed by

$$\xi_{\rm trans}(i) = \sum_{n=1}^{M/L} \gamma_n \bar{c}_n^{\,i} \tag{79}$$

where $\gamma_n = \boldsymbol{\lambda}^{\rm H}\mathbf{g}_n\mathbf{g}_n^{\rm H}[\mathbf{x}(0) - \mathbf{x}(\infty)]$. Note that the transient component $\xi_{\rm trans}(i) \to 0$ as $i \to \infty$. By using the existing expression for the transient component of the full-rank receiver described by $\xi_{\rm trans}^{\rm full-rank}(i) = \sum_{n=1}^{M} \gamma_n c_n^i$ [19], we can establish conditions for which the transient component of the INT receiver defined in (79) can vanish faster, i.e., the INT scheme converges faster. If the INT scheme reduces the eigenvalue spread of its covariance matrix, we have, for the ith iteration, the following expression:

$$\sum_{n=1}^{M/L} \gamma_n \bar{c}_n^{\,i} < \sum_{n=1}^{M} \gamma_n c_n^i. \tag{80}$$

The above condition states that the transient component of the reduced-rank INT scheme has fewer decreasing modes and vanishes before that of the full-rank structure. To verify (80), we studied the eigenvalue spread of the covariance matrices of the INT and the full-rank schemes in an extensive set of scenarios. In all situations, the experiments indicate an increase in the convergence speed as well as that the INT can reduce the eigenvalue spread of the full-rank scheme.

Appendix C
Convergence Speed of the INT Scheme With RLS Algorithms

Here, we evaluate the convergence speed of the proposed INT receiver scheme through the MSE analysis of RLS algorithms. By using a similar analysis to [31, Ch. 13, pp. 573–579] and replacing the expectation operator with time averages, let us express weight error vector of the reduced-rank INT least squares solution

$$\mathbf{e}_w(i) = \mathbf{w}(i) - \mathbf{w}_{\rm opt} = \hat{\bar{\mathbf{R}}}^{-1}(i)\sum_{l=1}^{i}\mathbf{r}(l)e_o^*(l). \tag{81}$$

Using the definition for the weight error-correlation matrix $\mathbf{K}(i) = E[\mathbf{e}_w(i)\mathbf{e}_w^{\rm H}(i)]$ [31], we have

$$\mathbf{K}(i) = E\left[\hat{\bar{\mathbf{R}}}^{-1}(i)\sum_{l=1}^{i}\sum_{j=1}^{i}\mathbf{r}(l)e_o^*(l)e_o(j)\mathbf{r}^{\rm H}(j)\hat{\bar{\mathbf{R}}}^{-1}(i)\right]. \tag{82}$$

Assuming that $e_o(i)$ is taken from a zero-mean Gaussian process with variance σ^2, we have $E[e_o(l)e_o^*(j)] = \begin{cases}\sigma^2, & l = j \\ 0, & l \neq j\end{cases}$ and

$$\begin{aligned}\mathbf{K}(i) &= \sigma^2 E\left[\hat{\bar{\mathbf{R}}}^{-1}(i)\sum_{l=1}^{i}\sum_{j=1}^{i}\mathbf{r}(l)\mathbf{r}^{\rm H}(j)\hat{\bar{\mathbf{R}}}^{-1}(i)\right] \\ &= \sigma^2 E\left[\hat{\bar{\mathbf{R}}}^{-1}(i)\right]. \end{aligned} \tag{83}$$

By invoking the independence theory and using the fact that the estimate of the covariance matrix given by $\hat{\bar{\mathbf{R}}}^{-1}(i)$ is described by a complex Wishart distribution [31, Sec. 13.6], the expectation of the time-averaged estimate $\hat{\bar{\mathbf{R}}}^{-1}(i)$ is exactly

$$E[\hat{\bar{\mathbf{R}}}^{-1}(i)] = \frac{1}{i - M/L - 1}\bar{\mathbf{R}}^{-1}, \qquad i > M/L + 1 \tag{84}$$

where $\bar{\mathbf{R}}^{-1}$ is the theoretical reduced-rank covariance matrix, and thus

$$\mathbf{K}(i) = \frac{\sigma^2\bar{\mathbf{R}}^{-1}}{i - M/L - 1}, \qquad i > M/L + 1. \tag{85}$$

Using the expression that describes the excess MSE in (58), we get

$$\xi_{\rm exc}(i) = {\rm tr}\left[\bar{\mathbf{R}}\mathbf{K}(i)\right] = \frac{\sigma^2 M/L}{i - M/L - 1}, \qquad i > M/L + 1. \tag{86}$$

The above result shows that the learning curve of the RLS algorithm with the proposed reduced-rank structure converges in about $2M/L$ iterations, in contrast to the RLS with the full-rank scheme that requires about $2M$ iterations [31]. This means that the proposed scheme converges L times faster than the full-rank approach with RLS techniques. Another observation from (86) is that, as i increases, the excess MSE tends to zero (for $\lambda = 1$), and it is independent from the eigenvalue spread of $\hat{\bar{\mathbf{R}}}^{-1}(i)$.

REFERENCES

[1] M. L. Honig and H. V. Poor, "Adaptive interference suppression," in *Wireless Communications: Signal Processing Perspectives*, H. V. Poor and G. W. Wornell, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1998, ch. 2, pp. 64–128.

[2] M. L. Honig, S. L. Miller, M. J. Shensa, and L. B. Milstein, "Performance of adaptive linear interference suppression in the presence of dynamic fading," *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 635–645, Apr. 2001.

[3] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178–3188, Dec. 1994.

[4] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 4, pp. 685–697, May 1994.

[5] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 944–960, Jul. 1995.

[6] Z. Xu and M. K. Tsatsanis, "Blind adaptive algorithms for minimum variance CDMA receivers," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 180–194, Jan. 2001.

[7] A. Klein, G. Kaleh, and P. Baier, "Zero forcing and minimum mean- square-error equalization for multiuser detection in code-division multiple-access channels," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 276–287, May 1996.

[8] S. Buzzi and H. V. Poor, "Channel estimation and multiuser detection in long-code DS/CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1476–1487, Aug. 2001.

[9] Z. Xu and M. Tsatsanis, "Blind channel estimation for long code multiuser CDMA systems," *IEEE Trans. Signal Process.*, vol. 48, no. 4, pp. 988–1001, Apr. 2000.

[10] P. Liu and Z. Xu, "Joint performance study of channel estimation and multiuser detection for uplink long-code CDMA systems," *EURASIP J. Wireless Commun. Netw.—Special Issue Innovative Signal Transmission Detection Techniques Next Generation Cellular CDMA System*, vol. 2004, no. 1, pp. 98–112, Aug. 2004.

[11] L. Li, A. M. Tulino, and S. Verdu, "Design of reduced-rank MMSE multiuser detectors using random matrix methods," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 986–1008, Jun. 2004.

[12] A. M. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceler," *IEEE Trans. Signal Process.*, vol. 39, no. 1, pp. 76–84, Jan. 1991.

[13] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 677–690, Mar. 1998.

[14] Y. Song and S. Roy, "Blind adaptive reduced-rank detection for DS-CDMA signals in multipath channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1960–1970, Nov. 1999.

[15] R. Singh and L. B. Milstein, "Interference suppression for DS/CDMA," *IEEE Trans. Commun.*, vol. 47, no. 3, pp. 446–453, Mar. 1999.

[16] R. Singh and L. B. Milstein, "Adaptive interference suppression for DS/CDMA," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1902–1905, Nov. 2002.

[17] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2943–2959, Nov. 1998.

[18] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 986–994, Jun. 2002.

[19] S. L. Miller, "Training analysis of adaptive interference suppression for direct-sequence code-division multiple-access systems," *IEEE Trans. Commun.*, vol. 44, no. 4, pp. 488–495, Apr. 1996.

[20] Y. Neuvo, C. Y. Dong, and S. K. Mitra, "Interpolated finite impulse response filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-32, no. 3, pp. 563–570, Jun. 1984.

[21] T. Saramaki, Y. Neuvo, and S. K. Mitra, "Design of computationally efficient interpolated FIR filters," *IEEE Trans. Circuits Syst.*, vol. 35, no. 1, pp. 70–88, Jan. 1988.

[22] A. Abousaada, T. Abousnasr, and W. Steenaart, "An echo tail canceller based on adaptive interpolated FIR filtering," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 39, no. 7, pp. 409–416, Jul. 1992.

[23] L. S. Resende, C. A. F. Rocha, and M. G. Bellanger, "A linearly constrained approach to the interpolated FIR filtering problem," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2000, pp. 392–395.

[24] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank interference suppression for DS-CDMA based on interpolated FIR filters," *IEEE Commun. Lett.*, vol. 9, no. 3, pp. 213–215, Mar. 2005.

[25] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank interference suppression for DS-CDMA using adaptive interpolated FIR filters with adaptive interpolators," in *Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Barcelona, Spain, Sep. 2004, pp. 150–154.

[26] R. C. de Lamare and R. Sampaio-Neto, "Blind adaptive reduced-rank CDMA receivers based on interpolated FIR filters with adaptive interpolators in multipath channels," in *Proc. IEEE GLOBECOM Conf.*, Dallas, TX, Dec. 2004, pp. 2493–2497.

[27] X. G. Doukopoulos and G. V. Moustakides, "Blind channel estimation for downlink CDMA systems," in *Proc. IEEE Int. Conf. Commun.*, 2003, pp. 2416–2420.

[28] X. G. Doukopoulos and G. V. Moustakides, "Power techniques for blind adaptive channel estimation in CDMA systems," in *Proc. IEEE Global Commun. Conf.*, 2003, pp. 2330–2334.

[29] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.

[30] D. S. Watkins, *Fundamentals of Matrix Computations*, 2nd ed. Hoboken, NJ: Wiley, 2002.

[31] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[32] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[33] T. S. Rappaport, *Wireless Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

Rodrigo C. de Lamare received the Diploma in electronic engineering from the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, in 1998 and the M.Sc. and Ph.D. degrees in electrical engineering from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), in 2001 and 2004, respectively.

From January 2004 to June 2005, he was a Postdoctoral Fellow with the Center for Studies in Telecommunications, PUC-Rio, and from July 2005 to January 2006, he worked as a Postdoctoral Fellow with the Signal Processing Laboratory, UFRJ. Since January 2006, he has been with the Communications Research Group, Department of Electronics, University of York, York, U.K., where he is currently a Lecturer in Communications Engineering. His research interests are in communications and signal processing.

Raimundo Sampaio-Neto received the Diploma and M.Sc. degrees in electrical engineering from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, in 1975 and 1978, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California (USC), Los Angeles, in 1983.

From 1978 to 1979, he was an Assistant Professor with PUC-Rio, and from 1979 to 1983, he was a Doctoral Student and a Research Assistant with the Department of Electrical Engineering, USC, with a Fellowship from CAPES. From November 1983 to June 1984, he was a Postdoctoral Fellow with the Communication Sciences Institute, Department of Electrical Engineering, USC, and was a member of the technical staff of Axiomatic Corporation, Los Angeles. He is currently a Researcher with the Center for Studies in Telecommunications, PUC-Rio, and has been an Associate Professor with the Department of Electrical Engineering, PUC-Rio, since July 1984. In 1991, he was a Visiting Professor with the Department of Electrical Engineering, USC. He has participated in various projects and has consulted for several private companies and government agencies. His areas of interest include communication-systems theory, digital transmission, satellite communications, and multiuser detection.

Prof. Sampaio-Neto was a Coorganizer of the Session on Recent Results for the IEEE Workshop on Information Theory, 1992, Salvador. He has also served as Technical Program Cochairman for IEEE Global Telecommunications Conference (Globecom'99) held in Rio de Janeiro in December 1999 and as a member of the technical program committees of several national and international conferences. He served two consecutive terms with the Board of Directors of the Brazilian Communications Society, where he is currently a member of its Advisory Council and Associate Editor of the *Journal of the Brazilian Communication Society*.