Adaptive Reduced-Rank Equalization Algorithms Based on Alternating Optimization Design Techniques for Multi-Antenna Systems

Rodrigo C. de Lamare and Raimundo Sampaio-Neto

***Abstract*—This paper presents a novel adaptive reduced-rank multi-input multi-output (MIMO) equalization scheme and algorithms based on alternating optimization design techniques for MIMO spatial multiplexing systems. The proposed reduced-rank equalization structure consists of a joint iterative optimization of two equalization stages, namely, a transformation matrix that performs dimensionality reduction and a reduced-rank estimator that retrieves the desired transmitted symbol. The proposed reduced-rank architecture is incorporated into an equalization structure that allows both decision feedback and linear schemes for mitigating the inter-antenna and inter-symbol interference. We develop alternating least squares (LS) expressions for the design of the transformation matrix and the reduced-rank estimator along with computationally efficient alternating recursive least squares (RLS) adaptive estimation algorithms. We then present an algorithm for automatically adjusting the model order of the proposed scheme. An analysis of the LS algorithms is carried out along with sufficient conditions for convergence and a proof of convergence of the proposed algorithms to the reduced-rank Wiener filter. Simulations show that the proposed equalization algorithms outperform the existing reduced-rank and full-rank algorithms, while requiring a comparable computational cost.**

***Index Terms*—MIMO systems, equalization structures, parameter estimation, reduced-rank schemes.**

I. Introduction

THE high demand for performance and capacity in wireless networks has led to the development of numerous signal processing and communications techniques for employing the resources efficiently. Recent results on information theory have shown that it is possible to achieve high spectral efficiency [1] and to make wireless links more reliable [2], [3] through the deployment of multiple antennas at both transmitter and receiver. In multi-input multi-output (MIMO) communications systems, the received signal is composed by the sum of several transmitted signals which share the propagation environment and are subject to multi-path propagation effects and noise at the receiver. The multipath channel originates inter-symbol interference (ISI), whereas the non-orthogonality among the signals transmitted gives rise to inter-antennas interference (IAI) at the receiver.

This work is partially funded by the Ministry of Defence (MoD), UK, Contract No. RT/COM/S/021. Part of this manuscript was presented at WCNC 2008. Dr. R. C. de Lamare is with the Communications Research Group, Department of Electronics, University of York, York Y010 5DD, United Kingdom and Prof. R. Sampaio-Neto is with CETUC/PUC-RIO, 22453-900, Rio de Janeiro, Brazil. E-mails: rcdl500@ohm.york.ac.uk and raimundo@cetuc.puc-rio.br

In order to mitigate the effects of ISI and IAI that reduce the performance and the capacity of MIMO systems the designer has to construct a MIMO equalizer. The optimal MIMO equalizer known as the maximum likelihood sequence estimation (MLSE) receiver was originally developed in the context of multiuser detection in [4]. However, the exponential complexity of the optimal MIMO equalizer makes its implementation costly for multipath channels with severe ISI and MIMO systems with many antennas. In practice, designers often prefer the deployment of low-complexity MIMO receivers such as the linear [5], [6], the successive interference cancellation-based VBLAST [7] and decision feedback equalizers (DFE) [8]-[14]. The DFE schemes [8]-[14] can achieve significantly better performance than linear ones due to the interference cancellation capabilities of the feedback section. These receivers require the estimation of the coefficients used for combining the received data and extracting the desired transmitted symbols. A challenging problem in MIMO systems [15] is encountered when the number of elements in the equalizer or the number of antenna pairs is large, which is key to future applications [16]-[18]. In these situations, an estimation algorithm requires substantial training for the MIMO equalizer and a large number of received symbols to reach its steady-state behavior.

There are many algorithms for designing MIMO equalizers, which possess different trade-offs between performance and complexity [19]. In this regard, least squares (LS)-based algorithms are often the preferred choice with respect to convergence performance. However, when the number of filter elements in the equalizer is large, an adaptive LS-type algorithm requires a large number of samples to reach its steady-state behavior and may encounter problems in tracking the desired signal. Reduced-rank techniques [20]-[34] are powerful and effective approaches in low-sample support situations and in problems with large filters. These algorithms can exploit the low-rank nature of signals that are found in MIMO communications [36] in order to achieve faster convergence speed, increased robustness against interference and better tracking performance than full-rank techniques. By projecting the input data onto a low-rank subspace associated with the signals of interest, reduced-rank methods can eliminate the interference that lies in the noise subspace and perform denoising [20]-[34]. Prior work on reduced-rank estimators for MIMO systems is extremely limited and relatively unexplored, being the work of Sun *et al.* [25] one of the few existing ones in the area. A comprehensive study of reduced-rank equalization algorithms for MIMO systems has not been considered so far. It is well known that the optimal reduced-rank approach is based on the eigen-decomposition (EVD) of the known input data covariance matrix $\boldsymbol{R}$ [20]. However, this covariance matrix must be estimated. The approach taken to estimate $\boldsymbol{R}$ and perform dimensionality reduction is of central importance and plays a key role in the performance of the system. Numerous reduced-rank strategies have been proposed in the last two decades. The first methods were based on the EVD of time-averaged estimates of $\boldsymbol{R}$ [20], in which the dimensionality reduction is carried out by a transformation matrix formed by appropriately selected eigenvectors computed with the EVD. A more recent and elegant approach to the problem was taken with the advent of the multistage Wiener filter (MSWF) [22], which was later extended to adaptive versions in [23], [24], and MIMO applications [25]. Another related method is the auxiliary vector filtering (AVF) algorithm [26]-[28], which can outperform the MSWF. A key limitation with prior art is the deficient exchange of information between the dimensionality reduction task and the subsequent reduced-rank estimation.

In this work, we propose adaptive reduced-rank MIMO equalization algorithms based on alternating optimization design techniques for MIMO spatial multiplexing systems. The proposed reduced-rank equalization structure and algorithms consist of a joint iterative optimization that alternates between two equalization stages, namely, a transformation matrix that performs dimensionality reduction and a reduced-rank estimator that suppresses the IAI caused by the associated data streams and retrieves the desired transmitted symbol. The essence of the proposed scheme is to change the role of the equalization filters and promote the exchange of information between the dimensionality reduction and the reduced-rank estimation tasks in an alternated way. In order to estimate the coefficients of the proposed MIMO reduced-rank equalizers, we develop alternating least squares (LS) optimization algorithms and expressions for the joint design of the transformation matrix and the reduced-rank filter. We derive alternating recursive LS (RLS) adaptive algorithms for their computationally efficient implementation and present a complexity study of the proposed and existing algorithms. We also describe an algorithm for automatically adjusting the model order of the proposed reduced-rank MIMO equalization schemes. An analysis of the proposed LS optimization is conducted, in which sufficient conditions and proofs for the convergence of the proposed algorithms are derived. The performance of the proposed scheme is assessed via simulations for MIMO equalization applications. The main contributions of this work are summarized as follows:

1) A reduced-rank MIMO equalization scheme and a design approach for both decision feedback and linear structures;
2) Reduced-rank LS expressions and recursive algorithms for parameter estimation;
3) An algorithm for automatically adjusting the model order;
4) Analysis and convergence proofs of the proposed algorithms.
5) A study of MIMO reduced-rank equalization algorithms.

This paper is structured as follows: The MIMO system and signal model is described in Section II. The proposed adaptive MIMO reduced-rank equalization structure is introduced along with the problem statement in Section III. Section IV is devoted to the development of the LS estimators, the computationally efficient RLS algorithms and the model order selection algorithms. Section V presents an analysis and proofs of convergence of the proposed algorithms. Section VI discusses the simulation results and Section VII gives the conclusions of this work.

Notation: In this paper bold upper and lowercase letters represent matrices and vectors, respectively. $(.)^*$, $(.)^*H$, $(.)^{-1}$ and $(.)^T$ shall represent complex conjugate, complex conjugate transpose (Hermitian), inverse and transpose, respectively. $\mathrm{tr}(.)$ is the trace operator of a matrix. Reduced-rank vectors and matrices are given with the addition of a bar $(\bar{.})$ and estimated symbols are denoted by the addition of a hat $(\hat{.})$.

II. MIMO System and Signal Model

In this section we present MIMO communications system and signal model and describe its main components. The model in this section is intended for describing a general MIMO system in multipath channels. However, it can also serve as a model for broadband MIMO communications systems with guard intervals including those based on orthogonal frequency-division multiplexing (OFDM) [37], [38] and single-carried (SC) modulation with frequency-domain equalization [39].

Consider a MIMO system with N_T antennas at the transmitter and N_R antennas at the receiver in a spatial multiplexing configuration, as shown in Fig. 1. The system is mathematically equivalent to that in [9]. The signals are modulated and transmitted from N_T antennas over multipath channels whose propagation effects are modelled by finite impulse response (FIR) filters with L_p coefficients, and are received by N_R antennas. We assume that the channel can vary during each packet transmission and the receiver is perfectly synchronized with the main propagation path. At the receiver, a MIMO equalizer is used to mitigate IAI and ISI and retrieve the transmitted signals.

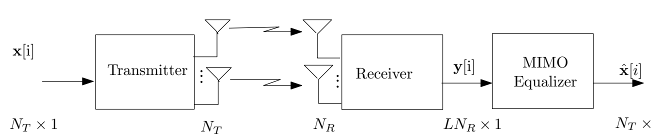

Fig. 1. MIMO system model.

The signals transmitted by the system at time instant i can be described by $\boldsymbol{x}[i] = [x_1[i] \ \ldots \ x_{N_T}[i]]^T$, where $x_j[i]$, $j = 1, \ \ldots \ N_T$ are independent and identically distributed symbols of unit variance. The demodulated signal received at the kth antenna and time instant i after applying a filter matched to the signal waveform and sampling at symbol rate is expressed by

$$y_k[i] = \sum_{j=1}^{N_T} \sum_{l=0}^{L_p-1} h_{j,k,l}[i] x_j[i-l] + n_k[i], \quad \text{for } k = 1, \ldots, N_R, \tag{1}$$

where $h_{j,k,l}[i]$ is the sampled impulse response between transmit antenna j and receive antenna k for path l, and $n_k[i]$ are samples of white Gaussian complex noise with zero mean and variance σ^2. By collecting the samples of the received signal and organizing them in a window of L symbols ($L \geq L_p$) for each antenna element, we obtain the $LN_R \times 1$ received vector

$$\boldsymbol{y}[i] = \boldsymbol{H}[i]\boldsymbol{x}_T[i] + \boldsymbol{n}[i], \quad (2)$$

where $\boldsymbol{y}[i] = \big[\boldsymbol{y}_1^T[i] \ \ldots \ \boldsymbol{y}_{N_R}^T[i]\big]^T$ contains the signals collected by the N_R antennas, the $L \times 1$ vector $\boldsymbol{y}_k[i] = \big[y_k[i] \ \ldots \ y_k[i-L+1]\big]^T$, for $k = 1, \ \ldots, \ N_R$, contains the signals collected by the kth antenna and are organized into a vector. The window size L must be chosen according to the prior knowledge about the delay spread of the multipath channel [45]. The $LN_R \times LN_T$ MIMO channel matrix $\boldsymbol{H}[i]$ is

$$\boldsymbol{H}[i] = \begin{bmatrix} \boldsymbol{H}_{1,1}[i] & \boldsymbol{H}_{1,2}[i] & \ldots & \boldsymbol{H}_{1,N_T}[i] \\ \boldsymbol{H}_{2,1}[i] & \boldsymbol{H}_{2,2}[i] & \ldots & \boldsymbol{H}_{2,N_T}[i] \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}_{N_R,1}[i] & \boldsymbol{H}_{N_R,2}[i] & \ldots & \boldsymbol{H}_{N_R,N_T}[i] \end{bmatrix}, \quad (3)$$

where the $L \times L$ matrix $\boldsymbol{H}_{j,k}[i]$ are Toeplitz matrices with the channel gains organized in a channel vector $\boldsymbol{h}_{j,k}[i] = [h_{j,k,1}[i] \ \ldots \ h_{j,k,L_p-1}]^T$ that is shifted down by one position from left to right for each column, and which describes the multi-path channel from antenna j to antenna k. The elements $h_{j,k,l}[i]$, for $l = 0, \ \ldots, \ L_p$, of $\boldsymbol{h}_{j,k}[i]$ are modelled as random variables and follow a specific propagation channel model [45], as will be detailed in the Section VI. The $LN_T \times 1$ vector $\boldsymbol{x}_T[i] = \big[\boldsymbol{x}_1^T[i] \ \ldots \ \boldsymbol{x}_{N_T}^T[i]\big]^T$ is composed by the data symbols transmitted from the N_T antennas at the transmitter with $\boldsymbol{x}_j[i] = [x_j[i] \ \ldots \ x_j[i-L+1]]^T$ being the ith transmitted block with dimensions $L \times 1$. The $LN_R \times 1$ vector $\boldsymbol{n}[i]$ is a complex Gaussian noise vector with zero mean and $E\big[\boldsymbol{n}[i]\boldsymbol{n}^H[i]\big] = \sigma^2\boldsymbol{I}$, where $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian transpose, respectively, and $E[\cdot]$ stands for expected value.

III. Proposed Adaptive Reduced-Rank MIMO DFE and Problem Formulation

We present the proposed reduced-rank MIMO equalization structure and state the main design problem of reduced-rank MIMO equalization structures. Both decision feedback (DF) and linear equalization structures can be devised by adjusting the dimensions of the filters and the use of feedback. We shall start with the description of the DF structure and then obtain the linear scheme as a particular case. In the proposed MIMO reduced-rank DF equalizer (DFE), the signal processing tasks are carried out in two stages, as illustrated in Fig. 2. The proposed scheme employs two sets of filters and stacks the decision and the input data vectors for joint processing. The decision feedback strategy adopted in this work is the parallel scheme reported in [9], [13], which firstly obtains the decision vector $\hat{\boldsymbol{x}}_{T,j}[i]$ with linear equalization and then employs $\hat{\boldsymbol{x}}_{T,j}[i]$ to cancel the interference caused by the interfering streams. A decision delay $\delta_{\rm dec}$ is assumed between the symbols transmitted and the $\hat{\boldsymbol{x}}_{T,j}[i]$ obtained after the decision block. The parallel strategy outperforms the successive one that uses a sequential procedure of equalization and interference cancellation [7], [8].

Let us consider the design of the proposed MIMO reduced-rank equalizer using the structure shown in Fig. 2. The $M \times 1$ input data vector $\mathbf{r}[i]$ to the proposed equalizer is obtained by stacking the $LN_R \times 1$ received vector $\boldsymbol{y}[i]$ and the $B(N_T - 1) \times 1$ vector of decisions $\hat{\boldsymbol{x}}_{T,j}[i]$ for stream j and is described by

$$\boldsymbol{r}_j[i] = \begin{bmatrix} \boldsymbol{y}[i] \\ \hat{\boldsymbol{x}}_{T,j}[i] \end{bmatrix}, \quad (4)$$

where $M = LN_R + B(N_T - 1)$ represents the number of samples for processing. The $B(N_T-1)\times 1$ vector of decisions $\hat{\boldsymbol{x}}_{T,j}[i] = [\hat{\boldsymbol{x}}_j[i] \ \ldots \ \hat{\boldsymbol{x}}_j[i-B+1]]^T$ for the jth stream takes into account B decision instants for the feedback and excludes the jth detected symbol to avoid cancelling the desired symbol. The $N_T \times 1$ vector of decisions is given by $\hat{\boldsymbol{x}}[i] = [\hat{x}_1[i] \ \ldots \ \hat{x}_{N_T}[i]]^T$, whereas the $N_T - 1 \times 1$ vector of decisions that excludes stream j and is employed to build $\hat{\boldsymbol{x}}_{T,j}[i]$ is given by $\hat{\boldsymbol{x}}_j[i] = [\hat{x}_1[i] \ \ldots \ \hat{x}_{j-1}[i] \ \hat{x}_{j+1}[i] \ \ldots \ \hat{x}_{N_T}[i]]^T$.

Let us now consider an $M \times D$ transformation matrix $\boldsymbol{S}_{D,j}[i]$ which carries out a dimensionality reduction on the received data $\boldsymbol{r}_j[i]$ and shall exploit the low-rank nature of the data transmitted over stream j as follows

$$\bar{\boldsymbol{r}}_j[i] = \boldsymbol{S}_{D,j}^H[i]\boldsymbol{r}_j[i], \quad j = 1, \ldots, N_T, \quad (5)$$

where D is the rank of the resulting equalization system.

The resulting projected received vector $\bar{\boldsymbol{r}}_j[i]$ is the input to an estimator represented by the $D \times 1$ vector $\bar{\boldsymbol{w}}_j[i] = [\bar{w}_{j,1}[i] \ \bar{w}_{j,2}[i] \ \ldots \ \bar{w}_{j,D}[i]]^T$. According to the schematic shown in Fig. 2, the output of the proposed MIMO reduced-rank DFE is obtained by linearly combining the coefficients of $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ for extracting the symbol transmitted from antenna j. Notice that all D-dimensional quantities have a "bar". The proposed MIMO reduced-rank DFE output is

$$\tilde{z}_j[i] = \bar{\boldsymbol{w}}_j^H[i]\boldsymbol{S}_{D,j}^H[i]\boldsymbol{r}_j[i] = \bar{\boldsymbol{w}}_j^H[i]\bar{\boldsymbol{r}}_j[i], \quad (6)$$

From the outputs $z_j[i]$ for $j = 1, \ 2, \ldots, N_T$, we construct the vector $\boldsymbol{z}[i] = \big[z_1[i] \ldots z_j[i] \ldots z_{N_T}[i]\big]^T$. The initial decisions for each data stream are obtained without resorting to the feedback and are computed as follows

$$\hat{x}_j[i] = Q\Big(\bar{\boldsymbol{w}}_j^H[i]\boldsymbol{S}_{D,j}^H[i]\begin{bmatrix} \boldsymbol{y}[i] \\ \boldsymbol{0} \end{bmatrix}\Big), \quad (7)$$

where $Q\big(\cdot\big)$ represents a decision device suitable for the constellation of interest (BPSK, QPSK or QAM) and the vector of decisions is constructed as $\hat{\boldsymbol{x}}[i] = \big[\hat{x}_1[i] \ldots \hat{x}_j[i] \ldots \hat{x}_{N_T}[i]\big]^T$ and used to construct $\hat{\boldsymbol{x}}_{T,j}[i]$ and $\boldsymbol{r}_j[i]$ as in (4). The detected symbols $\hat{\boldsymbol{x}}^{(f)}[i]$ of the proposed reduced-rank MIMO DFE after the IAI and ISI cancellation are obtained by

$$\hat{\boldsymbol{x}}^{(f)}[i] = Q\big(\boldsymbol{z}[i]\big) = Q\Bigg(\begin{bmatrix} \bar{\boldsymbol{w}}_1^H[i]\boldsymbol{S}_{D,1}^H[i]\boldsymbol{r}_1[i] \\ \vdots \\ \bar{\boldsymbol{w}}_{N_T}^H[i]\boldsymbol{S}_{D,N_T}^H[i]\boldsymbol{r}_{N_T}[i] \end{bmatrix}\Bigg). \quad (8)$$

The feedback employs $B(N_T - 1)$ connections for cancelling the IAI and the other $N_T - 1$ data streams and the ISI from the

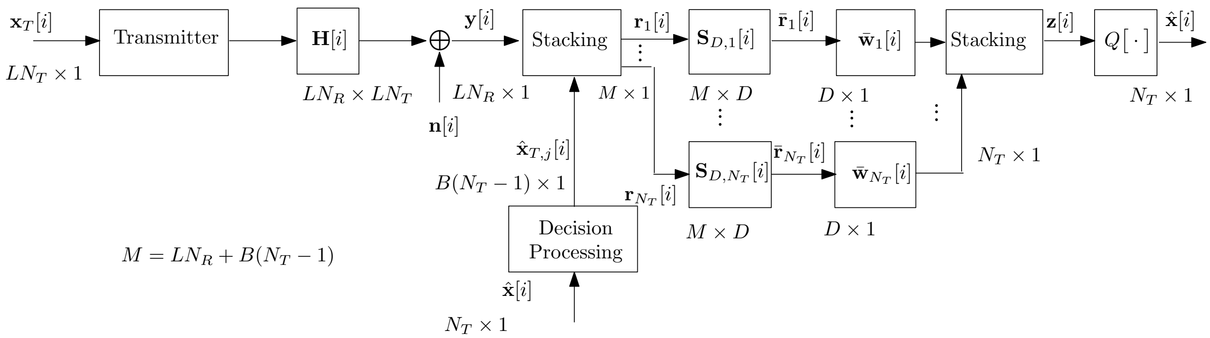

Fig. 2. Proposed MIMO reduced-rank decision feedback equalization structure.

adjacent symbols. A reduced-rank MIMO linear equalizer is obtained by neglecting the feedback with decision processing of the structure in Fig. 2.

The previous development suggests that the key aspect and problem to be solved in the design of reduced-rank MIMO equalization schemes is the cost-effective computation of the estimators $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$. The transformation matrix $\boldsymbol{S}_{D,j}[i]$ plays the most important role since it carries out the dimensionality reduction, which profoundly affects the performance of the remaining estimators and the MIMO equalizers. Methods based on the EVD [20], the MSWF [23] and the AVF [26]-[28] were reported for the design of $\boldsymbol{S}_{D,j}[i]$, however, they did not consider jointly the design of $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ via alternating optimization recursions. In the next section, we present the reduced-rank least squares (LS) algorithms and their recursive versions for the design of $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ used in the proposed MIMO equalization structure.

IV. Proposed Reduced-Rank Least Squares Design and Adaptive Algorithms

In this section, we present a joint iterative exponentially weighted reduced-rank LS estimator design of the parameters $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ of the proposed MIMO reduced-rank DFE. We then derive computationally efficient algorithms for computing the proposed LS estimator in a recursive way and automatically adjusting the model order. The deficient exchange of information between the dimensionality reduction task and the reduced-rank estimation verified in previously reported algorithms [22]-[28] is addressed by the alternated procedure that updates $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$. Specifically, the expression of $\boldsymbol{S}_{D,j}[i]$ is a function of $\bar{\boldsymbol{w}}_j[i]$ and vice versa, and this allows the coefficients to be computed via an alternating procedure with exchange of information in both ways (from $\boldsymbol{S}_{D,j}[i]$ to $\bar{\boldsymbol{w}}_j[i]$ and the other way around). Our studies and numerical results indicate that this approach is more effective than the MSWF [23] and the AVF [28] algorithms. In addition, the rank reduction is based on the joint and iterative LS minimization which has been found superior to the Krylov subspace, as evidenced in the numerical results. This allows the proposed method to outperform the MSWF and the AVF. We have opted for the use of one cycle (or iteration) per time instant in order to keep the complexity low. We also detail the computational complexity of the proposed and existing algorithms in terms of arithmetic operations.

A. Reduced-Rank Least Squares Estimator Design

In order to design $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$, we describe a joint iterative reduced-rank LS optimization algorithm. Consider the exponentially-weighted LS expressions for the estimators $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ via the cost function

$$\mathcal{C}_j(\boldsymbol{S}_{D,j}[i],\ \bar{\boldsymbol{w}}_j[i]) = \sum_{l=1}^{i} \lambda^{i-l} \big| x_j[l] - \bar{\boldsymbol{w}}_j^H[i] \boldsymbol{S}_{D,j}^H[i] \boldsymbol{r}_j[l] \big|^2, \tag{9}$$

where $0 < \lambda \leq 1$ is the forgetting factor.

The proposed exponentially-weighted LS design corresponds to solving the following optimization problem

$$\left\{ \boldsymbol{S}_{D,j}^{\rm opt},\ \bar{\boldsymbol{w}}_j^{\rm opt} \right\} = \arg \min_{\boldsymbol{S}_{D,j}[i], \boldsymbol{w}_j[i]} \mathcal{C}_j(\boldsymbol{S}_{D,j}[i],\ \bar{\boldsymbol{w}}_j[i]) \tag{10}$$

In order to solve the problem in (10), the proposed strategy is to fix a set of parameters, find the other set of parameters that minimize (9) and alternate this procedure between the two sets $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$. By minimizing (9) with respect to $\boldsymbol{S}_{D,j}[i]$, we obtain

$$\boldsymbol{S}_{D,j}[i] = \boldsymbol{R}_j^{-1}[i] \boldsymbol{P}_{D,j}[i] \boldsymbol{R}_{\bar{\boldsymbol{w}}_j}^{\dagger}[i-1], \tag{11}$$

where the $M \times D$ matrix $\boldsymbol{P}_{D,j}[i] = \sum_{l=1}^{i} \lambda^{i-l} x_j^*[l] \boldsymbol{r}_j[l] \bar{\boldsymbol{w}}_j^H[i-1]$, $\boldsymbol{R}_j[i] = \sum_{l=1}^{i} \lambda^{i-l} \boldsymbol{r}_j[l] \boldsymbol{r}_j^H[l]$, $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse and the $D \times D$ matrix $\boldsymbol{R}_{\bar{w}_j}[i-1] = \bar{\boldsymbol{w}}_j[i-1] \bar{\boldsymbol{w}}_j^H[i-1]$. Since $\boldsymbol{R}_{\bar{w}_j}[i-1]$ is a rank-1 matrix, we need to either compute the pseudo-inverse or introduce a regularization term in the recursion $\boldsymbol{R}_{\bar{w}_j}[i-1] = \sum_{l=1}^{i-1} \lambda^{i-l} \bar{\boldsymbol{w}}_j[l] \bar{\boldsymbol{w}}_j^H[l]$. We have opted for using the latter with the initial regularization factor $\boldsymbol{R}_{\bar{w}_j}[0] = \delta \boldsymbol{I}$ for numerical and simplicity reasons.

By minimizing (9) with respect to $\bar{\boldsymbol{w}}_j[i]$, the reduced-rank estimator becomes

$$\bar{\boldsymbol{w}}_j[i] = \bar{\boldsymbol{R}}_j^{-1}[i] \bar{\boldsymbol{p}}_j[i], \tag{12}$$

where $\bar{\boldsymbol{p}}_j[i] = \boldsymbol{S}_{D,j}^H[i] \sum_{l=1}^{i} \lambda^{i-l} x_j^*[l] \boldsymbol{r}_j[l] = \sum_{l=1}^{i} \lambda^{i-l} x_j^*[l] \bar{\boldsymbol{r}}_j[l]]$, and the $D \times D$ reduced-rank correlation matrix is described by $\bar{\boldsymbol{R}}_j[i] = \boldsymbol{S}_{D,j}^H[i] \sum_{l=1}^{i} \lambda^{i-l} \boldsymbol{r}_j[l] \boldsymbol{r}_j^H[l] \boldsymbol{S}_{D,j}[i]$.

The equation with the associated sum of error squares (SES) is obtained by substituting the expressions in (11) and (12) into the cost function (9), and is given by

$$\begin{aligned}\text{SES} = \sigma_{x_j}^2 - \bar{\boldsymbol{w}}_j^H[i]\boldsymbol{S}_{D,j}^H[i]\boldsymbol{p}[i] - \boldsymbol{p}^H[i]\boldsymbol{S}_{D,j}[i]\bar{\boldsymbol{w}}_j[i] \\ + \bar{\boldsymbol{w}}^H[i]\boldsymbol{S}_{D,j}^H[i]\boldsymbol{R}_j[i]\boldsymbol{S}_{D,j}[i]\bar{\boldsymbol{w}}_j[i],\end{aligned} \quad (13)$$

where $\sigma_{x_j}^2 = \sum_{l=1}^{i} \lambda^{i-l}|x_j[l]|^2$. Note that the expressions in (11) and (12) are not closed-form solutions for $\bar{\boldsymbol{w}}_j[i]$ and $\boldsymbol{S}_{D,j}[i]$ since they depend on each other and, thus, they have to be alternated with an initial guess to obtain a solution. The key strategy lies in the joint optimization of the estimators. The rank D must be set by the designer to ensure appropriate performance. The computational complexity of calculating (11) and (12) is cubic with the number of elements in the estimators, namely, M and D, respectively. In what follows, we introduce efficient RLS algorithms for computing the estimators with a quadratic cost.

B. Reduced-Rank RLS Algorithms

In this part, we present a recursive approach for efficiently computing the LS expressions developed in the previous subsection. Specifically, we develop reduced-rank RLS algorithms for computing $\bar{\boldsymbol{w}}_j[i]$ and $\boldsymbol{S}_{D,j}[i]$. Unlike conventional (full-rank) RLS algorithms that require the calculation of one estimator for the MIMO DFE, the proposed reduced-rank RLS technique jointly and iteratively computes the transformation matrix $\boldsymbol{S}_{D,j}[i]$ and the reduced-rank estimator $\bar{\boldsymbol{w}}_j[i]$. In order to start the derivation of the proposed algorithms, let us define

$$\begin{aligned}\boldsymbol{P}_j[i] &\triangleq \boldsymbol{R}_j^{-1}[i], \\ \boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-1] &\triangleq \boldsymbol{R}_{\bar{\boldsymbol{w}}_j}^{-1}[i-1], \\ \boldsymbol{P}_{D,j}[i] &\triangleq \lambda \boldsymbol{P}_{D,j}[i-1] + x_j^*[i]\boldsymbol{r}_j[i]\bar{\boldsymbol{w}}^H[i-1].\end{aligned} \quad (14)$$

Rewriting the expression in (11), we arrive at

$$\begin{aligned}\boldsymbol{S}_{D,j}[i] &= \boldsymbol{R}_j^{-1}[i]\boldsymbol{P}_{D,j}[i]\boldsymbol{R}_{\bar{\boldsymbol{w}}_j}^{-1}[i-1] \\ &= \boldsymbol{P}_j[i]\boldsymbol{P}_{D,j}[i]\boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-1] \\ &= \boldsymbol{S}_{D,j}[i-1] + \boldsymbol{k}_j[i]\big(x_j^*[i]\boldsymbol{t}_j^H[i-1] \\ &\quad - \boldsymbol{r}_j^H[i]\boldsymbol{S}_{D,j}[i-1]\big),\end{aligned} \quad (15)$$

where the $D \times 1$ vector $\boldsymbol{t}_j[i-1] = \boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-1]\bar{\boldsymbol{w}}_j[i-1]$, the $M \times 1$ Kalman gain vector is

$$\boldsymbol{k}_j[i] = \frac{\lambda^{-1}\boldsymbol{P}_j[i-1]\boldsymbol{r}_j[i]}{1+\lambda^{-1}\boldsymbol{r}_j^H[i]\boldsymbol{P}_j[i-1]\boldsymbol{r}_j[i]}, \quad (16)$$

the update for the $M \times M$ matrix $\boldsymbol{P}_j[i]$ employs the matrix inversion lemma [19]

$$\boldsymbol{P}_j[i] = \lambda^{-1}\boldsymbol{P}_j[i-1] - \lambda^{-1}\boldsymbol{k}_j[i]\boldsymbol{r}_j^H[i]\boldsymbol{P}_j[i-1], \quad (17)$$

and the $D \times 1$ vector $\boldsymbol{t}_j[i-1]$ is updated as

$$\boldsymbol{t}_j[i-1] = \frac{\lambda^{-1}\boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-1]\bar{\boldsymbol{w}}_j[i-1]}{1+\lambda^{-1}\bar{\boldsymbol{w}}_j^H[i-1]\boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-1]\bar{\boldsymbol{w}}_j[i-1]}. \quad (18)$$

The matrix inversion lemma is then used to update the $D \times D$ matrix $\boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-1]$ as described by

$$\boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-1] = \lambda^{-1}\boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-2] - \lambda^{-1}\boldsymbol{t}_j[i-1]\bar{\boldsymbol{w}}_j^H[i-2]\boldsymbol{Q}_{\bar{\boldsymbol{w}}_j}[i-2]. \quad (19)$$

Equations (14)-(19) constitute the part of the proposed reduced-rank RLS algorithms for computing $\boldsymbol{S}_{D,j}[i]$.

In order to develop the second part of the algorithm that estimates $\bar{\boldsymbol{w}}_j[i]$, let us consider the expression in (12) with its associated quantities, i.e., the $D \times D$ matrix $\bar{\boldsymbol{R}}_j[i] = \sum_{l=1}^{i} \lambda^{i-l}\bar{\boldsymbol{r}}_j[l]\bar{\boldsymbol{r}}_j^H[l]$ and the $D \times 1$ vector $\bar{\boldsymbol{p}}_j[i] = \sum_{l=1}^{i} \lambda^{i-l}x_j^*[l]\bar{\boldsymbol{r}}_j[l]$.

Let us now define $\bar{\boldsymbol{\Phi}}_j[i] = \boldsymbol{R}_j^{-1}[i]$ and rewrite $\bar{\boldsymbol{p}}_j[i]$ as $\bar{\boldsymbol{p}}_j[i] = \lambda\bar{\boldsymbol{p}}_j[i-1] + x_j^*[i]\bar{\boldsymbol{r}}_j[i]$. We can then rewrite (12) as follows

$$\begin{aligned}\bar{\boldsymbol{w}}_j[i] &= \bar{\boldsymbol{\Phi}}_j[i]\bar{\boldsymbol{p}}_j[i] \\ &= \bar{\boldsymbol{w}}_j[i-1] - \bar{\boldsymbol{k}}_j[i]\bar{\boldsymbol{r}}_j^H[i]\bar{\boldsymbol{w}}_j[i-1] + \bar{\boldsymbol{k}}_j[i]x_j^*[i] \\ &= \bar{\boldsymbol{w}}_j[i-1] + \bar{\boldsymbol{k}}_j[i]\big(x_j^*[i] - \bar{\boldsymbol{r}}_j^H[i]\bar{\boldsymbol{w}}_j[i-1]\big).\end{aligned} \quad (20)$$

By defining $\xi_j[i] = x_j[i] - \bar{\boldsymbol{w}}_j^H[i-1]\bar{\boldsymbol{r}}_j[i]$ we arrive at the proposed RLS algorithm for computing $\bar{\boldsymbol{w}}_j[i]$

$$\bar{\boldsymbol{w}}_j[i] = \bar{\boldsymbol{w}}_j[i-1] + \bar{\boldsymbol{k}}_j[i]\xi_j^*[i], \quad (21)$$

where the $D \times 1$ Kalman gain vector is given by

$$\bar{\boldsymbol{k}}_j[i] = \frac{\lambda^{-1}\bar{\boldsymbol{\Phi}}_j[i-1]\bar{\boldsymbol{r}}_j[i]}{1+\lambda^{-1}\bar{\boldsymbol{r}}_j^H[i]\bar{\boldsymbol{\Phi}}_j[i-1]\bar{\boldsymbol{r}}_j[i]}, \quad (22)$$

and the update for the matrix inverse $\bar{\boldsymbol{\Phi}}[i]$ employs the matrix inversion lemma [19]

$$\bar{\boldsymbol{\Phi}}_j[i] = \lambda^{-1}\bar{\boldsymbol{\Phi}}_j[i-1] - \lambda^{-1}\bar{\boldsymbol{k}}_j[i]\bar{\boldsymbol{r}}_j^H[i]\bar{\boldsymbol{\Phi}}_j[i-1]. \quad (23)$$

Equations (21)-(23) constitute the second part of the proposed algorithm that computes $\bar{\boldsymbol{w}}_j[i]$. The computational complexity of the proposed RLS algorithms is $O(D^2)$ for the estimation of $\bar{\boldsymbol{w}}_j[i]$ and $O(M^2)$ for the estimation of $\boldsymbol{S}_{D,j}[i]$. Since $D << M$ for moderate to large L, N_R, N_T and B, as will be explained in the next section, the overall complexity is in the same order of the conventional full-rank RLS algorithm ($O(M^2)$) [19].

C. Model-Order Selection Algorithm

The performance of the LS and RLS algorithms described in the previous subsection depends on the model order or the rank D. This motivates the development of methods to automatically adjust D using an LS cost function as a mechanism to control the selection. Prior methods for model order selection which use MSWF-based algorithms [23] or AVF-based recursions [28] have considered projection techniques [23] and cross-validation [28] approaches. Here, we focus on an approach that jointly determines D based on an LS criterion computed by the estimators $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$, where the superscript D denotes the rank used for the adaptation. The methods considered here (the proposed and existing ones [23], [28]) are the most suitable for model-order adaptation in time-varying channels. Other techniques such as the Akaike information criterion-based and the minimum description length do not lend themselves to time-varying situations and are computationally complex [19].

The key quantities to be updated are the transformation matrix $\boldsymbol{S}_{D,j}[i]$, the reduced-rank estimator $\bar{\boldsymbol{w}}_j[i]$, and the inverse of the reduced-rank covariance matrix $\bar{\boldsymbol{P}}_j[i]$ (for the proposed

RLS algorithm). Specifically, we allow the dimensions of $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ to vary from $D_{\min}$ and $D_{\max}$, which are the minimum and maximum ranks allowed, respectively. It is important to note that only one recursion to obtain $\bar{\boldsymbol{P}}_j[i]$ is computed with $D_{\max}$ in order to keep the complexity low. Once $\bar{\boldsymbol{P}}_j[i]$ is obtained, we perform a search for the best D for $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ that require sub-matrices of $\bar{\boldsymbol{P}}_j[i]$ for their computation. The transformation matrix $\boldsymbol{S}_{D,j}[i]$ and the reduced-rank estimator $\bar{\boldsymbol{w}}_j[i]$ employed with this algorithm are illustrated by

$$\boldsymbol{S}_{D,j}[i] = \begin{bmatrix} s_{1,1,j}[i] & \ldots & s_{1,D_{\min},j}[i] & \ldots & s_{1,D_{\max},j}[i] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{M,1,j}[i] & \ldots & s_{M,D_{\min},j}[i] & \ldots & s_{M,D_{\max},j}[i] \end{bmatrix}$$

and

$$\bar{\boldsymbol{w}}_{D,j}[i] = \left[\; w_{1,j}[i] \; w_{2,j}[i] \; \ldots \; w_{D_{\min},j}[i] \; \ldots \; w_{D_{\max},j}[i] \;\right]^T \tag{24}$$

The method for automatically selecting D of the algorithm is based on the exponentially weighted *a posteriori* least-squares type cost function:

$$\mathcal{C}_j(\boldsymbol{S}_{D,j}[i], \bar{\boldsymbol{w}}_{d,j}[i]) = \sum_{l=1}^{i} \lambda^{i-l} \left| x_j[l] - \bar{\boldsymbol{w}}_{d,j}^H[i] \boldsymbol{S}_{D,j}^H[i] \boldsymbol{r}_j[l] \right|^2. \tag{25}$$

For each time interval i, we select the rank $D_{\rm opt}[i]$ which minimizes $\mathcal{C}_j(\boldsymbol{S}_{D,j}[i], \bar{\boldsymbol{w}}_{D,j}[i])$ and the exponential weighting factor λ is required as the optimal rank varies as a function of the data record. The transformation matrix $\boldsymbol{S}_{D,j}[i]$ and the reduced-rank estimator $\bar{\boldsymbol{w}}_{d,j}[i]$ are updated along with $\bar{\boldsymbol{P}}[i]$ for the maximum allowed rank $D_{\max}$ and then the proposed rank adaptation algorithm determines the the best model order for each time instant i using the cost function in (25). The proposed model-order selection algorithm is given by

$$D_{j,{\rm opt}}[i] = \arg \min_{D_{\min} \leq d \leq D_{\max}} \mathcal{C}_j(\boldsymbol{S}_{d,j}[i], \bar{\boldsymbol{w}}_{d,j}[i]), \tag{26}$$

where d is an integer, $D_{\min}$ and $D_{\max}$ are the minimum and maximum ranks allowed for the estimators, respectively. A small rank may provide faster adaptation during the initial stages of the estimation procedure, whereas a large rank usually yields a better steady-state performance. Our studies indicate that the range for which the rank D of the proposed algorithms have a positive impact on the performance of the algorithms is limited. Specifically, we have found that even for large systems ($N_R = N_T = 20, 30, 40, 50, 60$) the rank does not scale with the system size and remains small. The typical range of values remains between $D_{\min} = 3$ and $D_{\max} = 8$ for the system sizes examined ($N_R = N_T = 20, 30, 40, 50, 60$). This is an important aspect of the proposed algorithms because it keeps the complexity low (comparable to a standard RLS algorithm). For the scenarios considered in what follows, we set $D_{\min} = 3$ and $D_{\max} = 8$. In the simulations section, we will illustrate how the proposed model-order selection algorithm performs.

D. Computational Complexity

In this subsection, we illustrate the computational complexity requirements of the proposed RLS algorithms and compare them with those of existing algorithms. We also provide the computation complexity of the proposed and existing model-order selection algorithms. The computational complexity of the algorithms is expressed in terms of additions and multiplications, as depicted in Table I. For the proposed reduced-rank RLS algorithm the complexity is quadratic with $M = LN_R + B(N_T - 1)$ and D. This amounts to a complexity slightly higher than that observed for the full-rank RLS algorithm, provided D is significantly smaller than M, and significantly less than the cost of the MSWF-RLS [23] and the AVF [28] algorithms. The complexity of the proposed model-order selection algorithm is given in Table II.

TABLE I
COMPUTATIONAL COMPLEXITY OF ALGORITHMS.

Algorithm	Additions	Multiplications
Full-rank [19]	$2M^2$ $+M+1$	$3M^2$ $+5M$
Proposed	$2M^2$ $-M+4D^2$ $+MD+D+3$	$3M^2$ $+3M+6D^2$ $+MD+8D$
MSWF [23]	DM^2 $+6D^2-8D+2$ $+M^2$	DM^2 $+2DM+3D$ $+M^2+2$
AVF [28]	DM^2+2M-1 $+5D(M-1)+1$ $+3(DM-1)^2$	$4DM^2$ $+4DM+4M$ $+4D+2$

In order to illustrate the main trends and requirements in terms of complexity of the proposed and existing algorithms, we show in Fig. 3 the complexity against the number of input samples M for the parameters $D = 5$, $N_T = N_R$, $L = 8$ and $B = 2$. The curves indicate that the proposed reduced-rank RLS algorithm has a complexity significantly lower than the MSWF-RLS algorithm [23] and the AVF [28], whereas it remains at the same level of the full-rank RLS algorithm.

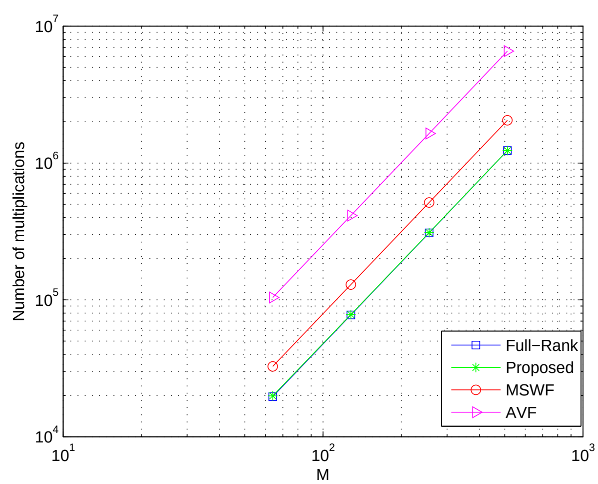

Fig. 3. Complexity in terms of multiplications against number of input samples (M) with $D = 5$, $N_T = N_R$, $L = 8$, $B = 2$.

The computational complexity of the model-order selection algorithms including the proposed and the existing techniques is shown in Table II. We can notice that the proposed model-

TABLE II
COMPUTATIONAL COMPLEXITY OF MODEL-ORDER SELECTION ALGORITHMS.

Algorithm	Additions	Multiplications
Proposed	$2(D_{\max} - D_{\min}) + 1$	$-$
Projection with Stopping Rule [23]	$2(2M-1)\times$ $(D_{\max} - D_{\min}) + 1$	$(M^2+M+1)\times$ $(D_{\max} - D_{\min} + 1)$
CV [28]	$(2M-1)\times$ $(2(D_{\max} - D_{\min}) + 1)$	$(D_{\max} - D_{\min} + 1)\times$ $M+1$

order selection algorithm is significantly less complex than the existing methods based on projection with stopping rule [23] and the CV approach [28]. Specifically, the proposed algorithm that uses extended filters only requires $2(D_{\max} - D_{\min})$ additions, as depicted in the first row of Table II. To this cost we must add the operations required by the proposed RLS algorithm, whose complexity is shown in the second row of Table I using $D_{\max}$. The complexities of the MSWF and the AVF algorithms are detailed in the third and fourth rows of Table I. For their operation with model-order selection algorithms, a designer must add their complexities in Table I to the complexity of the model-order selection algorithms of interest in Table II.

V. Analysis of The Proposed Algorithms

In this section, we conduct an analysis of the proposed algorithms that compute the estimators $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ of the proposed scheme. We first highlight the alternating optimization nature of the proposed algorithms and make use of recent convergence results for this class of algorithms [40], [41]. In particular, we present a set of sufficient conditions under which the proposed algorithms converge to the optimal estimators. This is corroborated by our numerical studies that verify that the method is insensitive to different initializations (except for the case when $\boldsymbol{S}_{D,j}[i]$ is a null matrix which annihilates the received signal) and that it converges to the same point of minimum. We establish the global convergence of the proposed algorithm via induction and show that that the sequence of estimators $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ produces a sequence of outputs that is bounded and converges to the reduced-rank Wiener filter [20],[21].

A. Sufficient Conditions for Convergence

In order to develop the analysis and proofs, we need to define a metric space and the Hausdorff distance that will be used extensively. A metric space is an ordered pair $(\mathcal{M}, d)$ where $\mathcal{M}$ is a non-empty set and d is a metric on $\mathcal{M}$, i.e., a function $d : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ such that for any x, y, z, and $\mathcal{M}$ we have:
a) $d(x, y) \geq 0$.
b) $d(x, y) = 0 \; iff \; x = y$.
c) $d(x, y) = d(y, x)$.
d) $d(x, z) \leq d(x, y) + d(y, z$ (triangle inequality).
The Hausdorff distance measures how far two subsets of a metric space are from each other and is defined by

$$d_H(X,Y) = \max\{\sup_{x\in X} \inf_{y \in Y} d(x,y), \sup_{y\in Y} \inf_{x \in X} d(x,y)\} \quad (27)$$

The proposed LS and RLS algorithms can be stated as an alternating minimization strategy based on the sum of error squares (SES) defined in (13) and expressed as

$$\boldsymbol{S}_{D,j}[i] \in \arg \min_{\boldsymbol{S}_{D,j}^{\rm opt} \in \underline{\boldsymbol{S}}_{D,j}[i]} {\rm SES}(\boldsymbol{S}_{D,j}^{\rm opt}, \bar{\boldsymbol{w}}_j[i]) \quad (28)$$

$$\bar{\boldsymbol{w}}_j[i] \in \arg \min_{\bar{\boldsymbol{w}}_j^{\rm opt} \in \underline{\bar{\boldsymbol{w}}}_j[i]} {\rm SES}(\boldsymbol{S}_{D,j}[i], \bar{\boldsymbol{w}}_j^{\rm opt}), \quad (29)$$

where $\boldsymbol{S}_{D,j}^{\rm opt}$ and $\bar{\boldsymbol{w}}_j^{\rm opt}$ correspond to the optimal values of $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$, respectively, and the sequences of compact sets $\{\underline{\boldsymbol{S}}_{D,j}[i]\}_{i\geq 0}$ and $\{\underline{\bar{\boldsymbol{w}}}_j[i]\}_{i\geq 0}$ converge to the sets $\underline{\boldsymbol{S}}_{D,{\rm opt}}$ and $\underline{\bar{\boldsymbol{w}}}_{j,{\rm opt}}$, respectively.

Although we are not given the sets $\underline{\boldsymbol{S}}_{D,{\rm opt}}$ and $\underline{\bar{\boldsymbol{w}}}_{j,{\rm opt}}$ directly, we observe the sequence of compact sets $\{\underline{\boldsymbol{S}}_{D,j}[i]\}_{i\geq 0}$ and $\{\underline{\bar{\boldsymbol{w}}}_j[i]\}_{i\geq 0}$. The goal of the proposed algorithms is to find a sequence of $\boldsymbol{S}_{D,j}[i]$ and $\bar{\boldsymbol{w}}_j[i]$ such that

$$\lim_{i\rightarrow\infty} {\rm SES}(\boldsymbol{S}_{D,j}[i], \bar{\boldsymbol{w}}_j[i]) = {\rm SES}(\boldsymbol{S}_{D,j}^{\rm opt}, \bar{\boldsymbol{w}}_j^{\rm opt}) \quad (30)$$

In order to present a set of sufficient conditions under which the proposed algorithms converge, we need to so-called ”three-point” and ”four-point” properties [40], [41]. Let us assume that there is a function $f : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ such that the following conditions are satisfied:

1) Three-point property ($\boldsymbol{S}_{D,j}^{\rm opt}$, $\tilde{\boldsymbol{S}}_{D,j}$, $\bar{\boldsymbol{w}}_j^{\rm opt}$): for all $i \geq 1$, $\boldsymbol{S}_{D,j}^{\rm opt} \in \underline{\boldsymbol{S}}_{D,j}[i]$, $\bar{\boldsymbol{w}}_j^{\rm opt} \in \underline{\bar{\boldsymbol{w}}}_j[i]$, and $\tilde{\boldsymbol{S}}_{D,j} \in \arg\min_{\bar{\boldsymbol{w}}_j^{\rm opt} \in \underline{\bar{\boldsymbol{w}}}_j[i]} {\rm SES}(\boldsymbol{S}_{D,j}^{\rm opt}, \bar{\boldsymbol{w}}_j^{\rm opt})$

$$f(\boldsymbol{S}_{D,j}^{\rm opt}, \tilde{\boldsymbol{S}}_{D,j}) + {\rm SES}(\tilde{\boldsymbol{S}}_{D,j}, \bar{\boldsymbol{w}}_j^{\rm opt}) \leq {\rm SES}(\boldsymbol{S}_{D,j}^{\rm opt}, \bar{\boldsymbol{w}}_j^{\rm opt}). \quad (31)$$

2) Four-point property ($\boldsymbol{S}_{D,j}^{\rm opt}$, $\bar{\boldsymbol{w}}_j^{\rm opt}$, $\tilde{\boldsymbol{S}}_{D,j}$, $\tilde{\bar{\boldsymbol{w}}}_j^{\rm opt}$): for all $i \geq 1$, $\boldsymbol{S}_{D,j}^{\rm opt}$, $\tilde{\boldsymbol{S}}_{D,j} \in \underline{\boldsymbol{S}}_{D,j}[i]$, $\bar{\boldsymbol{w}}_j^{\rm opt} \in \underline{\bar{\boldsymbol{w}}}_j[i]$, and $\tilde{\bar{\boldsymbol{w}}}_{D,j} \in \arg\min_{\bar{\boldsymbol{w}}_j^{\rm opt} \in \underline{\bar{\boldsymbol{w}}}_j[i]} {\rm SES}(\tilde{\boldsymbol{S}}_{D,j}, \bar{\boldsymbol{w}}_j^{\rm opt})$

$${\rm SES}(\boldsymbol{S}_{D,j}^{\rm opt}, \tilde{\bar{\boldsymbol{w}}}_j) \leq {\rm SES}(\boldsymbol{S}_{D,j}^{\rm opt}, \bar{\boldsymbol{w}}_j^{\rm opt}) + f(\boldsymbol{S}_{D,j}^{\rm opt}, \tilde{\boldsymbol{S}}_{D,j}). \quad (32)$$

Theorem: Let $\{(\underline{\boldsymbol{S}}_{D,j}[i], \underline{\bar{\boldsymbol{w}}}_j[i])\}_{i\geq 0}$, $\underline{\boldsymbol{S}}_{D,j}^{\rm opt}$, $\underline{\bar{\boldsymbol{w}}}_j^{\rm opt}$ be compact subsets of the compact metric space $(\mathcal{M}, d)$ such that

$$\underline{\boldsymbol{S}}_{D,j}[i] \xrightarrow{d_H} \underline{\boldsymbol{S}}_{D,j}^{\rm opt}, \qquad \underline{\bar{\boldsymbol{w}}}_j[i] \xrightarrow{d_H} \underline{\bar{\boldsymbol{w}}}_j^{\rm opt} \quad (33)$$

and let ${\rm SES} : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ be a continuous function.

Now let conditions **1**) and **2**) hold. Then, for the proposed algorithms we have

$$\lim_{i\rightarrow\infty} {\rm SES}(\boldsymbol{S}_{D,j}[i], \bar{\boldsymbol{w}}_j[i]) = {\rm SES}(\boldsymbol{S}_{D,j}^{\rm opt}, \bar{\boldsymbol{w}}_j^{\rm opt}) \quad (34)$$

A general proof of this theorem is detailed in [40], [41].

B. Convergence to the Optimal Reduced-Rank Estimator

In this subsection, we show that the proposed reduced-rank algorithm converges globally and exponentially to the optimal reduced-rank estimator [20],[21]. We assume that $1 \geq \lambda \gg 0$ (equal or close to one) , the desired product of the optimal solutions, i.e., $\boldsymbol{w}_j^{\rm opt} = \boldsymbol{S}_{D,j}^{\rm opt} \bar{\boldsymbol{w}}_j^{\rm opt}$ is known and given by $\boldsymbol{R}_j^{-1/2}[i] \big(\boldsymbol{R}_j^{-1/2}[i] \boldsymbol{p}_j[i]\big)_{1:D}$ [19],[21], where $\boldsymbol{R}_j^{-1/2}[i]$ is the square root of the input data covariance matrix and the subscript $1:D$ denotes truncation of the subspace.

In order to proceed with our proof, let us rewrite the expressions in (11) and (12) for time instant 0 as follows

$$\boldsymbol{R}_j[0]\boldsymbol{S}_{D,j}[0]\boldsymbol{R}_{\bar{w}_j}[0] = \boldsymbol{P}_{D,j}[0] = \boldsymbol{p}_j[0]\bar{\boldsymbol{w}}_j^H[0], \quad (35)$$

$$\bar{\boldsymbol{R}}_j[0]\bar{\boldsymbol{w}}_j[1] = \boldsymbol{S}_{D,j}^H[0]\boldsymbol{R}_j[0]\boldsymbol{S}_{D,j}[0]\bar{\boldsymbol{w}}_j[1] = \bar{\boldsymbol{p}}_j[0], \quad (36)$$

Using (35) we can obtain the following relation

$$\boldsymbol{R}_{\bar{w}_j}[0] = \big(\boldsymbol{S}_{D,j}^H[0]\boldsymbol{R}_j^2[0]\boldsymbol{S}_{D,j}[0]\big)^{-1}\boldsymbol{S}_{D,j}^H[0]\boldsymbol{R}_j[0]\boldsymbol{p}_j[0]\bar{\boldsymbol{w}}_j^H[0], \quad (37)$$

Substituting the above result for $\boldsymbol{R}_{w_j}[0]$ into the expression in (35) we get a recursive expression for $\boldsymbol{S}_{D,j}[0]$

$$\boldsymbol{S}_{D,j}[0] = \boldsymbol{R}_j[0]^{-1}\boldsymbol{p}_j[0]\bar{\boldsymbol{w}}_j^H[0]\big(\boldsymbol{S}_{D,j}^H[0]\boldsymbol{R}_j[0]\boldsymbol{p}_j[0]\bar{\boldsymbol{w}}_j^H[0]\big)^{-1}\cdot \cdot\big(\boldsymbol{S}_{D,j}^H[0]\boldsymbol{R}_j^2[0]\boldsymbol{S}_{D,j}[0]\big)^{-1}, \quad (38)$$

Using (36) we can express $\bar{\boldsymbol{w}}_j[1]$ as

$$\bar{\boldsymbol{w}}_j[1] = \big(\boldsymbol{S}_{D,j}^H[0]\boldsymbol{R}_j[0]\boldsymbol{S}_{D,j}[0]\big)^{-1}\boldsymbol{S}_{D,j}^H[0]\boldsymbol{p}_j[0], \quad (39)$$

Employing the relation $\boldsymbol{w}_j[1] = \boldsymbol{S}_{D,j}[1]\bar{\boldsymbol{w}}_j[1]$, we obtain

$$\boldsymbol{w}_j[1] = \boldsymbol{R}_j[1]^{-1}\boldsymbol{p}[1]\bar{\boldsymbol{w}}_j^H[1]\big(\boldsymbol{S}_{D,j}^H[1]\boldsymbol{R}_j[1]\boldsymbol{p}_j[1]\bar{\boldsymbol{w}}_j^H[1]\big)^{-1}\cdot \cdot\big(\boldsymbol{S}_{D,j}^H[1]\boldsymbol{R}_j^2[1]\boldsymbol{S}_{D,j}[1]\big)^{-1}\big(\boldsymbol{S}_{D,j}^H[0]\boldsymbol{R}_j[0]\boldsymbol{S}_{D,j}[0]\big)^{-1}\boldsymbol{S}_{D,j}^H[0]\boldsymbol{p}_j[0] \quad (40)$$

More generally, we can express the proposed reduced-rank LS algorithm by the following recursion

$$\begin{aligned}\boldsymbol{w}_j[i] &= \boldsymbol{S}_{D,j}[i]\bar{\boldsymbol{w}}_j[i]\\ &= \boldsymbol{R}_j[i]^{-1}\boldsymbol{p}_j[i]\bar{\boldsymbol{w}}_j^H[i]\big(\boldsymbol{S}_{D,j}^H[i]\boldsymbol{R}_j[i]\boldsymbol{p}_j[i]\bar{\boldsymbol{w}}_j^H[i]\big)^{-1}\cdot\\ &\cdot\big(\boldsymbol{S}_{D,j}^H[i]\boldsymbol{R}^2[i]\boldsymbol{S}_{D,j}[i]\big)^{-1}\cdot\\ &\cdot\big(\boldsymbol{S}_{D,j}^H[i-1]\boldsymbol{R}_j[i-1]\boldsymbol{S}_{D,j}[i-1]\big)^{-1}\boldsymbol{S}_{D,j}^H[i-1]\boldsymbol{p}_j[i-1].\end{aligned} \quad (41)$$

Since the optimal reduced-rank filter can be described by the EVD of $\boldsymbol{R}_j^{-1/2}[i]\boldsymbol{p}_j[i]$ [20], [21], where $\boldsymbol{R}_j^{-1/2}[i]$ is the square root of the covariance matrix $\boldsymbol{R}_j[i]$ and $\boldsymbol{p}_j[i]$ is the cross-correlation vector, then we have

$$\boldsymbol{R}_j^{-1/2}[i]\boldsymbol{p}_j[i] = \boldsymbol{\Phi}_j\boldsymbol{\Lambda}_j\boldsymbol{\Phi}_j^H\boldsymbol{p}_j[i], \quad (42)$$

where $\boldsymbol{\Lambda}_j$ is a $M\times M$ diagonal matrix with the eigenvalues of $\boldsymbol{R}_j$ and $\boldsymbol{\Phi}_j$ is a $M\times M$ unitary matrix with the eigenvectors of $\boldsymbol{R}_j$. Assuming that there exists some $\boldsymbol{w}_j[0]$ such that the randomly selected $\boldsymbol{S}_{D,j}[0]$ can be written as [21]

$$\boldsymbol{S}_{D,j}[0] = \boldsymbol{R}_j^{-1/2}[i]\boldsymbol{\Phi}_j\boldsymbol{w}_j[0]. \quad (43)$$

Using (43) and (42) in (41), and manipulating the algebraic expressions, we can express (41) in a more compact way that is suitable for analysis, as given by

$$\boldsymbol{w}_j[i] = \boldsymbol{\Lambda}_j^2\boldsymbol{w}_j[i-1](\boldsymbol{w}_j^H[i-1]\boldsymbol{\Lambda}_j^2\boldsymbol{w}_j[i-1])^{-1}\boldsymbol{w}_j^H[i-1]\boldsymbol{w}_j[i-1]. \quad (44)$$

The above expression can be decomposed as follows

$$\boldsymbol{w}_j[i] = \boldsymbol{Q}_j[i]\ \boldsymbol{Q}_j[i-1]\ \ldots\ \boldsymbol{Q}_j[1]\ \boldsymbol{w}_j[0], \quad (45)$$

where

$$\boldsymbol{Q}_j[i] = \boldsymbol{\Lambda}_j^{2i}\boldsymbol{w}_j[0](\boldsymbol{w}_j^H[0]\boldsymbol{\Lambda}_j^{4i-2}\boldsymbol{w}_j[0])^{-1}\boldsymbol{w}_j^H[0]\boldsymbol{\Lambda}_j^{2i-2}. \quad (46)$$

At this point, we need to establish that the norm of $\boldsymbol{S}_{D,j}[i]$ for all i is both lower and upper bounded, i.e., $0 < ||\boldsymbol{S}_{D,j}[i]|| < \infty$ for all i, and that $\boldsymbol{w}_j[i] = \boldsymbol{S}_{D,j}[i]\bar{\boldsymbol{w}}_j[i]$ approaches $\boldsymbol{w}_{j,\mathrm{opt}}[i]$ exponentially as i increases. Due to the linear mapping, the boundedness of $\boldsymbol{S}_{D,j}[i]$ is equivalent to that of $\boldsymbol{w}_j[i]$. Therefore, we have upon convergence $\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i-1] = \boldsymbol{w}_j^H[i-1]\boldsymbol{w}_j[i-1]$. Since $||\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i-1]|| \leq ||\boldsymbol{w}_j[i-1]||||\boldsymbol{w}_j[i]||$ and $||\boldsymbol{w}_j^H[i-1]\boldsymbol{w}_j[i-1]|| = ||\boldsymbol{w}_j[i-1]||^2$, the relation $\boldsymbol{w}_j^H[i]\boldsymbol{w}_j[i-1] = \boldsymbol{w}_j^H[i-1]\boldsymbol{w}_j[i-1]$ implies $||\boldsymbol{w}_j[i]|| > ||\boldsymbol{w}_j[i-1]||$ and hence

$$||\boldsymbol{w}_j[\infty]|| \geq ||\boldsymbol{w}_j[i]|| \geq ||\boldsymbol{w}_j[0]|| \quad (47)$$

In order to show that the upper bound $||\boldsymbol{w}_j[\infty]||$ is finite, let us express the $M\times M$ matrix $\boldsymbol{Q}_j[i]$ as a function of the $M\times 1$ vector $\boldsymbol{w}_j[i] = \begin{bmatrix}\boldsymbol{w}_{j,1}[i]\\ \boldsymbol{w}_{j,2}[i]\end{bmatrix}$ and the $M\times M$ matrix $\boldsymbol{\Lambda} = \begin{bmatrix}\boldsymbol{\Lambda}_{j,1} & \\ & \boldsymbol{\Lambda}_{j,2}\end{bmatrix}$. Substituting the previous expressions of $\boldsymbol{w}_j[i]$ and $\boldsymbol{\Lambda}_j$ into $\boldsymbol{Q}_j[i]$ given in (46), we obtain

$$\begin{aligned}\boldsymbol{Q}_j[i] = &\begin{bmatrix}\boldsymbol{\Lambda}_{j,1}^{2i}\boldsymbol{w}_{j,1}[0]\\ \boldsymbol{\Lambda}_{j,2}^{2i}\boldsymbol{w}_{j,2}[0]\end{bmatrix}(\boldsymbol{w}_{j,1}^H[0]\boldsymbol{\Lambda}_{j,1}^{4i-2}\boldsymbol{w}_{j,1}[0]\\ &+\boldsymbol{w}_{j,2}^H[0]\boldsymbol{\Lambda}_2^{4i-2}\boldsymbol{w}_{j,2}[0])^{-1}\begin{bmatrix}\boldsymbol{w}_{j,1}^H[0]\boldsymbol{\Lambda}_{j,1}^{2i-2}\\ \boldsymbol{w}_{j,2}^H[0]\boldsymbol{\Lambda}_{j,2}^{2i-2}\end{bmatrix}.\end{aligned} \quad (48)$$

Using the matrix identity $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{A}^{-1}$ to the decomposed $\boldsymbol{Q}_j[i]$ in (48) and making i large, we get

$$\boldsymbol{Q}_j[i] = \mathrm{diag}\big(\underbrace{1\ldots1}_{D}\ \underbrace{0\ldots0}_{M-D}\big) + \mathrm{O}(\epsilon[i]). \quad (49)$$

where $\epsilon[i] = (\lambda_{r+1}/\lambda_r)^{2i}$ with λ_{r+1} and λ_r are the $(r+1)$th, the rth largest singular values of $\boldsymbol{R}_j^{-1/2}[i]\boldsymbol{p}_j[i]$ and $\mathrm{O}(\cdot)$ denotes the order of the argument. From (49), it follows that for some positive constant k, we have $||\boldsymbol{w}_j[i]|| \leq 1 + k\epsilon[i]$. From (45), we obtain

$$\begin{aligned}||\boldsymbol{w}_j[\infty]|| &\leq ||\boldsymbol{Q}_j[\infty]||\ldots||\boldsymbol{Q}_j[2]||\ ||\boldsymbol{Q}_j[1]||\ ||\boldsymbol{Q}_j[0]||\\ &\leq ||\boldsymbol{w}_j[0]||\prod_{i=1}^{\infty}(1+k\epsilon[i])\\ &= ||\boldsymbol{w}_j[0]||\exp\Big(\sum_{i=1}^{\infty} log(1+k\epsilon[i])\Big)\\ &\leq ||\boldsymbol{w}_j[0]||\exp\Big(\sum_{i=1}^{\infty} k\epsilon[i]\Big)\\ &= ||\boldsymbol{w}_j[0]||\exp\Big(\frac{k}{1-(\lambda_{r+1}/\lambda_r)^2}\Big)\end{aligned} \quad (50)$$

With the development above, the norm of $\boldsymbol{w}_j[i]$ is proven to be both lower and upper bounded. Once this is established, the expression in (41) converges for a sufficiently large i to the reduced-rank Wiener filter. This is verified by equating the terms of (44) which yields

$$\begin{aligned}\boldsymbol{w}_j[i] = \boldsymbol{R}_j[i]^{-1}\boldsymbol{p}_j[i]\bar{\boldsymbol{w}}_j^H[i]\big(\boldsymbol{S}_{D,j}^H[i]\boldsymbol{R}_j[i]\boldsymbol{p}_j[i]\bar{\boldsymbol{w}}_j^H[i]\big)^{-1}\cdot\\ \cdot\big(\boldsymbol{S}_{D,j}^H[i]\boldsymbol{R}_j^2[i]\boldsymbol{S}_{D,j}[i]\big)^{-1}\cdot\\ \cdot\big(\boldsymbol{S}_{D,j}^H[i-1]\boldsymbol{R}_j[i-1]\boldsymbol{S}_{D,j}[i-1]\big)^{-1}\boldsymbol{S}_{D,j}^H[i-1]\boldsymbol{p}_j[i-\\ = \boldsymbol{R}_j^{-1/2}[i]\boldsymbol{\Phi}_{j,1}\boldsymbol{\Lambda}_{j,1}\boldsymbol{\Phi}_{j,1}^H\boldsymbol{p}_j[i] + \mathrm{O}(\epsilon[i]),\end{aligned} \tag{51}$$

where $\boldsymbol{\Phi}_1$ is a $M \times D$ matrix with the D largest eigenvectors of $\boldsymbol{R}_j[i]$ and $\boldsymbol{\Lambda}_{j,1}$ is a $D \times D$ matrix with the largest eigenvalues of $\boldsymbol{R}_j[i]$.

VI. Simulation Results

In this section, we evaluate the bit error rate (BER) performance of the proposed MIMO equalization structure, algorithms and existing techniques, namely, the full-rank [9], the reduced-rank MSWF [23], and the AVF [28] techniques for the design of the receivers. For all simulations and the proposed reduced-rank RLS algorithm, we use the initial values $\bar{\boldsymbol{w}}_j[0] = \big[1, 0, \ldots, 0\big]$ and $\boldsymbol{S}_{D,j}[0] = [\boldsymbol{I}_D\ \boldsymbol{0}_{D\times(M-D)}]^T$. For the next experiments, we adopt an observation window of $L = 8$, the multipath channels (the channel vectors $\boldsymbol{h}_{j,k}[i] = [h_{j,k,1}[i]\ \ldots\ h_{j,k,L_p-1}]^T$) are modelled by FIR filters with L_p coefficients spaced by one symbol and the system employs QPSK modulation. The channel is time-varying over the transmitted packets, the profile follows the UMTS Vehicular A channel model [44] with $L_p = 5$ and the fading is given by Clarke's model [45]. We average the experiments over 1000 runs and define the signal-to-noise ratio (SNR) as $\mathrm{SNR} = 10\log_{10}\frac{N_T\sigma_x^2}{\sigma^2}$, where σ_x^2 is the variance of the transmitted symbols and σ^2 is the noise variance. The adaptive MIMO equalizers employ $N_T = 4$, $B = 4$, $L = 8$ and $N_R = 8$ in a spatial multiplexing configuration, leading to estimators at the receiver with $M = LN_R + B(N_T - 1) = 72$ coefficients. The adaptive RLS estimators of all methods are trained with 250 symbols, employ $\lambda = 0.998$ unless otherwise specified, and are then switched to decision-directed mode.

A. Convergence Performance and Impact of Rank

In the first experiment, we consider the BER performance versus the rank D with optimized parameters (forgetting factor $\lambda = 0.998$) for linear MIMO equalizers. The curves in Fig. 4 show that the best rank for the proposed scheme is $D = 4$, which is the closest among the analyzed algorithms to the optimal linear MMSE that assumes the knowledge of the channel and the noise variance. In addition, it should be remarked that our studies with systems with different sizes show that the optimal rank D does not vary significantly with the system size. It remains in a small range of values, which brings considerably faster convergence speed. However, It should also be remarked that the optimal rank D depends on the data record size and other parameters of the systems.

The BER convergence performance versus the number of received symbols for MIMO decision feedback equalizers with optimized but fixed ranks is shown in Fig. 5. The results show that the proposed scheme has a significantly faster

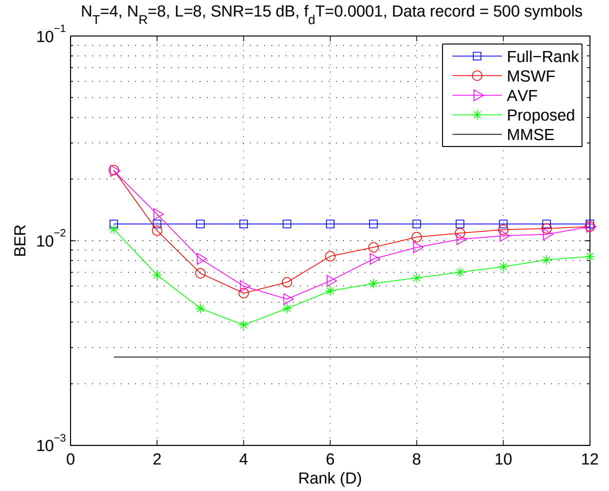

Fig. 4. BER performance versus rank (D) for linear MIMO equalizers.

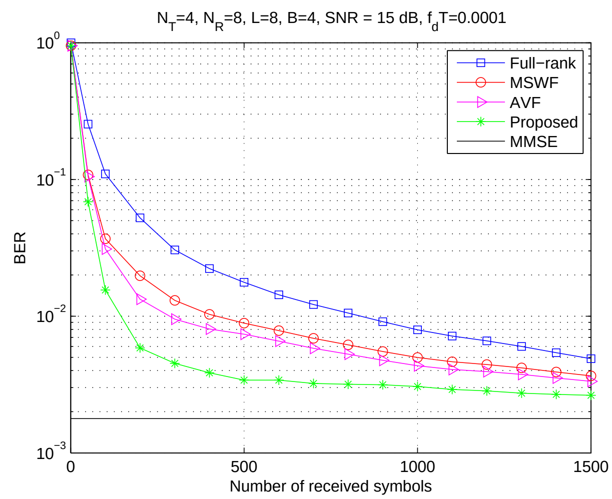

Fig. 5. BER performance versus number of received symbols.

convergence performance and obtains good gains over the best known schemes. The plots show that the proposed reduced-rank MIMO equalizer extends the dimensionality reduction and its benefits such as fast convergence and robustness against errors to the MIMO equalization task. The proposed RLS estimation algorithm has the best performance and is followed by the AVF, the MSWF, and the full-rank estimators. Note that the BER of the considered techniques will converge to the same values if the number of received symbols is very large and the channel is static.

B. Performance with Model Order Selection

As previously mentioned in Section IV, it is possible to further increase the convergence speed and enhance the tracking performance of the reduced-rank algorithms using an automatic model-order selection algorithm. In the next experiment, we consider the proposed reduced-rank structures and algorithms with linear and DF equalizers and compare their performance with fixed ranks and the proposed automatic model-order selection algorithm developed in Section IV.C. The results illustrated in Fig. 6 show that the proposed model-order selection algorithm can effectively speed up the convergence of the proposed reduced-rank RLS algorithm and ensure that it obtains an excellent tracking performance. In what follows, we will consider the proposed model-order selection algorithm in conjunction with the proposed reduced-rank RLS algorithm, and for a fair comparison we will equip the MSWF and the AVF algorithms with the rank adaptation techniques reported in [23] and [28], respectively.

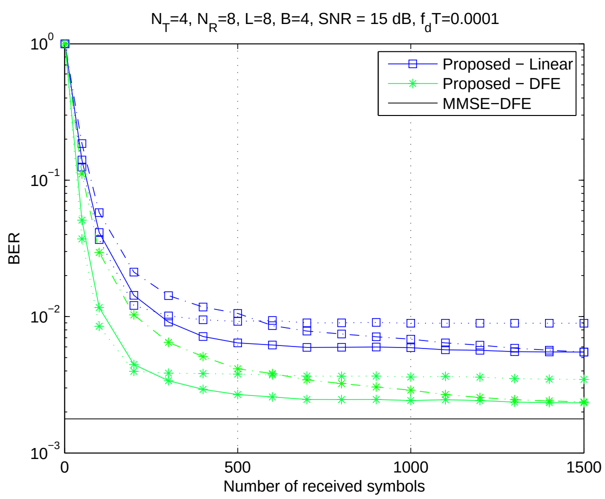

Fig. 6. BER performance versus number of received symbols for proposed estimation algorithms and structures. The performance of the proposed reduced-rank algorithms is shown for the proposed model-order selection algorithm (solid lines), for $D = 3$ (dotted lines) and for $D = 8$ (dash-dotted lines).

C. Performance for Various SNR and f_dT Values

The BER performance versus the signal-to-noise ratio (SNR) for MIMO decision feedback equalizers operating with the automatic model-order selection algorithms is shown in Fig. 7. The curves show a significant advantage of reduced-rank algorithms over the full-rank RLS algorithm. Specifically, the reduced-rank AVF and MSWF techniques obtain gains of up to 3 dB in SNR for the same BER over the full-rank algorithm, whereas the proposed reduced-rank RLS algorithm achieves a gain of up to 3 dB over the AVF, the second best reduced-rank algorithm. The main reasons for the differences in diversity order are the speed and the level of accuracy of the parameter estimation of the proposed and existing methods. If we increase the number of received symbols to a very large value then the diversity order attained by the different analyzed algorithms would be the same as verified in our studies.

In order to assess the performance of the reduced-rank algorithms for different fading rates, we consider an experiment where we measure the BER of the proposed and analyzed algorithms against the normalized fading rate f_dT in cycles per symbol, where f_d is the maximum Doppler frequency and T is the symbol rate. It should be noted that the forgetting factor λ was optimized for each value of f_dT in this experiment. In practice, a designer could employ a mechanism

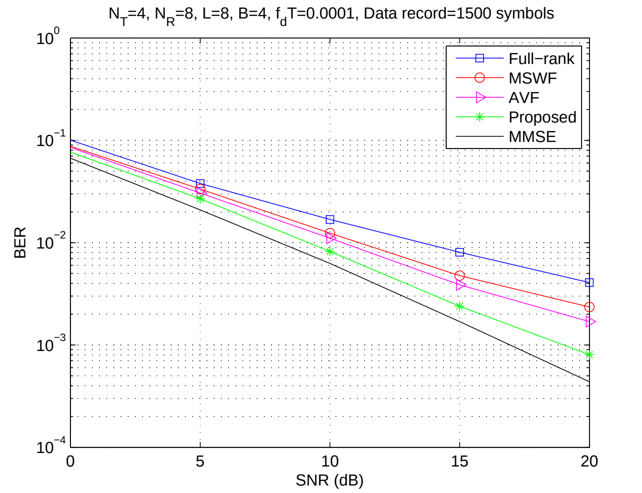

Fig. 7. BER performance versus SNR .

to automatically adjust λ. The results of this experiment are shown in Fig. 8, where the advantages of the reduced-rank algorithms and their superior performance in time-varying scenarios is verified again.

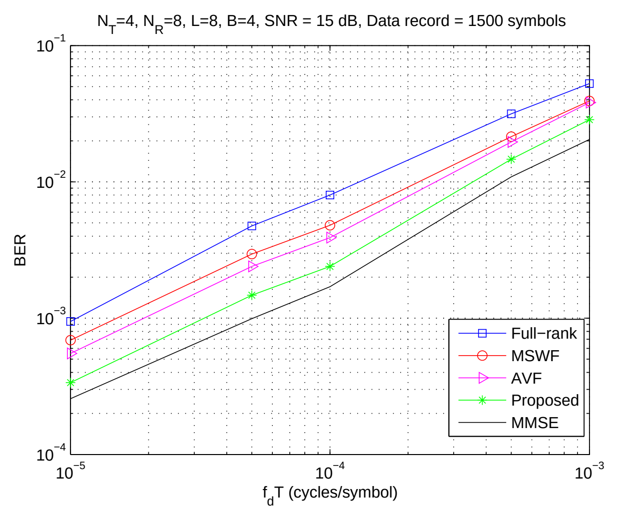

Fig. 8. BER performance versus the normalized fading rate f_dT .

D. Performance for MIMO-OFDM Systems

In the previous experiments, we considered the proposed MIMO equalization structure and algorithms for time-varying channels that dynamically change within a packet transmission, thereby requiring the adaptive equalization techniques so far described. At this point, it would be important to address two additional issues. The first is to account for the gains of reduced-rank techniques over full-rank methods when the order of the estimators changes. Another important aspect to be investigated is the applicability of the proposed reduced-rank techniques to broadband communications such as MIMO-OFDM systems [37], [38]. Even though in MIMO-OFDM systems the frequency selective channels are transformed into frequency flat channels, there is still the need to perform spatial equalization. We consider an experiment with a MIMO-OFDM system in which the data streams per sub-carrier are separated by MIMO linear equalizers equipped with full-rank and reduced-rank algorithms and the channels change at each OFDM block. The system has $N = 64$ sub-carriers and employs a cyclic prefix that corresponds to $C = 8$ symbols. The channel profile is identical to the model employed for the previous experiments and the fading is independent for each stream. The $N_R \times 1$ received data vector for the nth subcarrier is given by

$$\boldsymbol{r}_n[i] = \boldsymbol{H}_n[i]\boldsymbol{x}_n[i] + \boldsymbol{n}_n[i], \quad n = 1, 2, \ldots, N, \quad (52)$$

where the $N_R \times N_T$ channel matrix $\boldsymbol{H}_n[i]$ contains the channel frequency response gains at the nth tone, the $N_T \times 1$ data vector $\boldsymbol{x}_n[i]$ corresponds to the symbols transmitted by the N_T antennas over the nth subcarrier and the $N_R \times 1$ vector $\boldsymbol{n}_n[i]$ represents the noise vector at the nth tone.

We employ the proposed MIMO linear equalization scheme for spatial equalization on a per subcarrier-basis [37] for the OFDM symbols with the proposed and analyzed reduced-rank estimation algorithms. The BER is plotted against the number of antennas in a MIMO-OFDM system that has $N_R = N_T$. The results in Fig. 9 show that the advantages of reduced-rank algorithms are more pronounced for larger systems, in which the training requirements are more demanding in term of training data for full-rank RLS algorithms.

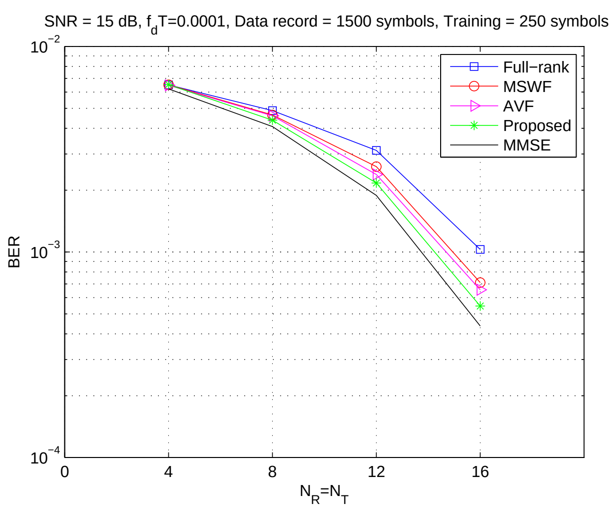

Fig. 9. BER performance versus the number of antennas .

The advantages of the reduced-rank estimators are due to the reduced amount of training and the relatively short data record (packet size). Therefore, for packets with relatively small size, the faster training of reduced-rank LS estimators will lead to superior BER to conventional full-rank LS estimators. As the length of the packets is increased, the advantages of reduced-rank estimators become less pronounced for training purposes and so become the BER advantages over full-rank estimators. In comparison with the MSWF and AVF reduced-rank schemes, the proposed scheme exploits the joint and iterative exchange of information between the transformation matrix and the reduced-rank estimators, which leads to better performance. The gains of the reduced-rank techniques over full-rank methods for MIMO-OFDM systems are less pronounced than those observed for narrowband MIMO systems with multipath channels. This is because the number of coefficients for estimation is significantly reduced. If we increase the number of antennas in MIMO-OFDM systems to a large value, then the gains of reduced-rank techniques become larger.

VII. Concluding Remarks

This paper has presented a study of reduced-rank equalization algorithms for MIMO systems. We have proposed an adaptive reduced-rank MIMO equalization scheme and algorithms based on joint iterative optimization of adaptive estimators. We have developed LS expressions and efficient RLS algorithms for the design of the proposed reduced-rank MIMO equalizers. A model-order selection algorithm for automatically adjusting the model order of the proposed algorithm has also been developed. An analysis of the convergence of the proposed algorithm has been carried out and proofs of global convergence of the algorithms have been established. Simulations for MIMO equalization applications have shown that the proposed schemes outperforms the state-of-the-art reduced-rank and the conventional estimation algorithms at a comparable computational complexity. Future work and extensions of the proposed scheme may consider strategies with iterative processing [13], [25], [46] with the aid of convolutional, Turbo and LDPC codes, detection structures which attain a higher diversity order [13], [14] and their theoretical analysis.

References

[1] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Person. Commun.*, vol. 6, pp. 311–335, Mar. 1998.

[2] S. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451-1458, Oct. 1998.

[3] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, “Space-time block codes from orthogonal designs,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[4] S. Verdú, “Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels”, *IEEE Transactions on Information Theory*, vol.IT-32, no. 1, pp. 85-96, Jan., 1986.

[5] A. Duel-Hallen, “Equalizers for Multiple Input Multiple Output Channels and PAM Systems with Cyclostationary Input Sequences,” *IEEE Journal on Selected Areas in Communications*, vol. 10, pp. 630-639, April, 1992.

[6] R. C. de Lamare and R. Sampaio-Neto, “Blind adaptive MIMO receivers for space-time block-coded DS-CDMA systems in multipath channels using the constant modulus criterion,” *IEEE Transactions on Communications*, vol. 58, no. 1, January 2010, pp. 21-27.

[7] G. D. Golden, G. J. Foschini, R. A. Valenzuela, P. W. Wolniansky, “Detection Algorithm and Initial Laboratory Results using the V-BLAST Space-Time Communication Architecture”, *Electronics Letters*, Vol. 35, No. 1, Jan. 7, 1999, pp. 14-15.

[8] G. Ginis and J. M. Cioffi, “On the relation between V-BLAST and the GDFE”, *IEEE Communications Letters*, vol. 5, issue 9, pp. 364-366, 2001.

[9] N. Al-Dhahir and A. H. Sayed, ”The finite-length multi-input multi-output MMSE-DFE,” *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2921-2936, Oct., 2000.

[10] C. Kominakis, C. Fragouli, A. H. Sayed and R. Wesel, ”Multi-input multi-output fading channel tracking and equalization using Kalman estimation,” *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1065-1076, May, 2002.

[11] R. C. de Lamare, R. Sampaio-Neto, “Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels”, *IEEE Communications Letters*, vol. 7, no. 2, Feb. 2003, pp. 73 - 75.

[12] R. C. de Lamare, R. Sampaio-Neto, A. Hjorungnes, “Joint iterative interference cancellation and parameter estimation for cdma systems”, *IEEE Communications Letters*, vol. 11, no. 12, December 2007, pp. 916 - 918.

[13] R. C. de Lamare and R. Sampaio-Neto, “Minimum Mean Squared Error Iterative Successive Parallel Arbitrated Decision Feedback Detectors for DS-CDMA Systems,” *IEEE Transactions on Communications*, vol. 56, no. 5, May 2008, pp. 778 - 789.

[14] J. H. Choi, H. J, Yu and Lee, Y. H., “Adaptive MIMO decision feedback equalization for receivers with time-varying channels,” IEEE Trans. Signal Process., vol. 53, no. 11, pp. 4295-4303, Nov., 2005.

[15] B. Hassibi, B. M. Hochwald, “How much training is needed in multiple-antenna wireless links?”, *IEEE Transactions on Information Theory*, vol. 49, no. 4, April 2003, pp. 951 - 963.

[16] C.-Y. Chiu, J.-B. Yan, R.D. Murch,“24-Port and 36-Port Antenna Cubes Suitable for MIMO Wireless Communications,” *IEEE Transactions on Antennas and Propagation*, vol.56, no.4, pp.1170-1176, April 2008.

[17] P. Li, R.D. Murch, “Multiple output selection-LAS algorithm in large MIMO systems,” *IEEE Communications Letters*, vol.14, no.5, pp.399-401, May 2010.

[18] S.K Mohammed, A. Zaki, A. Chockalingam, B. S. Rajan, “High-Rate Space-Time Coded Large-MIMO Systems: Low-Complexity Detection and Channel Estimation,” *IEEE Journal of Selected Topics in Signal Processing*, vol.3, no.6, pp.958-974, Dec. 2009.

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

[20] L. L. Scharf, “The SVD and reduced rank signal processing,” Signal Processing, vol. 25, no. 2, pp. 113–133, 1991.

[21] Y. Hua and M. Nikpour, “Computing the reduced rank Wiener filter by IQMD,” *IEEE Signal Processing Letters*, pp. 240-242, Vol. 6, Sept. 1999.

[22] J. S. Goldstein, I. S. Reed and L. L. Scharf, “A multistage representation of the Wiener filter based on orthogonal projections,” *IEEE Transactions on Information Theory*, vol. 44, November, 1998.

[23] M. L. Honig and J. S. Goldstein, “Adaptive reduced-rank interference suppression based on the multistage Wiener filter,” *IEEE Trans. on Communications*, vol. 50, no. 6, June 2002.

[24] R. C. de Lamare, M. Haardt and R. Sampaio-Neto, “Blind Adaptive Constrained Reduced-Rank Parameter Estimation based on Constant Modulus Design for CDMA Interference Suppression,” *IEEE Transactions on Signal Processing*, vol. 56., no. 6, June 2008.

[25] Y. Sun, V. Tripathi, and M. L. Honig, ”Adaptive, Iterative, Reduced-Rank (Turbo) Equalization”, *IEEE Trans. on Wireless Communications*, Vol. 4, No. 6, pp. 2789-2800, Nov. 2005.

[26] D. A. Pados, F. J. Lombardo and S. N. Batalama, “Auxiliary Vector Filters and Adaptive Steering for DS-CDMA Single-User Detection,” *IEEE Transactions on Vehicular Technology*, vol. 48, No. 6, November 1999.

[27] D. A. Pados, G. N. Karystinos, “An iterative algorithm for the computation of the MVDR filter,” *IEEE Trans. on Sig. Proc.*, vol. 49, No. 2, February, 2001.

[28] H. Qian and S.N. Batalama, “Data record-based criteria for the selection of an auxiliary vector estimator of the MMSE/MVDR filter”, *IEEE Trans. on Communications*, vol. 51, no. 10, Oct. 2003, pp. 1700 - 1708.

[29] R. C. de Lamare and R. Sampaio-Neto, “Adaptive reduced-rank MMSE filtering with interpolated FIR filters and adaptive interpolators”, *IEEE Signal Processing Letters*, vol. 12, no. 3, March, 2005.

[30] R. C. de Lamare and Raimundo Sampaio-Neto, “Reduced-rank Interference Suppression for DS-CDMA based on Interpolated FIR Filters”, *IEEE Communications Letters*, vol. 9, no. 3, March 2005.

[31] R. C. de Lamare and R. Sampaio-Neto, “Adaptive Interference Suppression for DS-CDMA Systems based on Interpolated FIR Filters with Adaptive Interpolators in Multipath Channels”, *IEEE Trans. Vehicular Technology*, Vol. 56, no. 6, September 2007.

[32] R. C. de Lamare and R. Sampaio-Neto, “Adaptive Reduced-Rank MMSE Parameter Estimation based on an Adaptive Diversity Combined Decimation and Interpolation Scheme,” *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, April 15-20, 2007, vol. 3, pp. III-1317-III-1320.

[33] R. C. de Lamare and R. Sampaio-Neto, Reduced-Rank Adaptive Filtering Based on Joint Iterative Optimization of Adaptive Filters, *IEEE Signal Processing Letters*, Vol. 14, no. 12, December 2007.

[34] R. C. de Lamare and R. Sampaio-Neto, “Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation, and Filtering,” *IEEE Transactions on Signal Processing*, vol. 57, no. 7, July 2009, pp. 2503 - 2514.

[35] R. C. de Lamare and R. Sampaio-Neto, “Reduced-Rank Space-Time Adaptive Interference Suppression With Joint Iterative Least Squares Algorithms for Spread-Spectrum Systems,” *IEEE Transactions on Vehicular Technology*, vol.59, no.3, March 2010, pp.1217-1228.

[36] D. Gesbert, H. Bolcskei, D. Gore, A. Paulraj, “Outdoor MIMO wireless channels: Models and performance prediction”, *IEEE Trans. on Communications*, December 2002.

[37] Y. Li, J. Winters, and Nelson Sollenberger, “MIMO-OFDM for Wireless Communications: Signal Detection with Enhanced Channel Estimation,” *IEEE Trans. on Communications*, Sept. 2002.

[38] G. L. Stuber, J. R. Barry, S. W. MacLaughlin, Y. Li, M. A. Ingram and T. G. Pratt ”Broadband MIMO-OFDM Wireless Communications”, Proceedings of the IEEE, vol. 92, no. 2, Feb 2004.

[39] D. Falconer, S, Lek Ariyavisitakul, A. Benyamin-Seeyar and B. Eidson, ”Frequency Domain Equalization for Single-Carrier Broadband Wireless Systems”, IEEE Communications Maganize, April 2002.

[40] I. Csiszár and G. Tusnády, “Information geometry and alternating minimization procedures,” *Statistics and Decisions*, Supplement Issue, no. l, pp. 205–237, 1984.

[41] U. Niesen, D. Shah, and G. W. Wornell, “Adaptive Alternating Minimization Algorithms”, *IEEE Trans. Inform. Theory*, vol. 55, no. 3, pp. 1423-1429, Mar. 2009.

[42] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, Md, 1996.

[43] D. Luenberger, *Linear and Nonlinear Programming*, 2nd Ed. Addison-Wesley, Inc., Reading, Massachusetts 1984.

[44] Third Generation Partnership Project (3GPP), specifications 25.101, 25.211-25.215, versions 5.x.x.

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

[46] X. Wang and H. V. Poor, “Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” *IEEE Trans. Commun.*, vol. 47, pp. 1046–1061, July 1999.