Adaptive and Iterative Multi-Branch MMSE Decision Feedback Detection Algorithms for Multi-Antenna Systems

Rodrigo C. de Lamare

Abstract—In this work, decision feedback (DF) detection algorithms based on multiple processing branches for multi-input multi-output (MIMO) spatial multiplexing systems are proposed. The proposed detector employs multiple cancellation branches with receive filters that are obtained from a common matrix inverse and achieves a performance close to the maximum likelihood detector (MLD). Constrained minimum mean-squared error (MMSE) receive filters designed with constraints on the shape and magnitude of the feedback filters for the multi-branch MMSE DF (MB-MMSE-DF) receivers are presented. An adaptive implementation of the proposed MB-MMSE-DF detector is developed along with a recursive least squares-type algorithm for estimating the parameters of the receive filters when the channel is time-varying. A soft-output version of the MB-MMSE-DF detector is also proposed as a component of an iterative detection and decoding receiver structure. A computational complexity analysis shows that the MB-MMSE-DF detector does not require a significant additional complexity over the conventional MMSE-DF detector, whereas a diversity analysis discusses the diversity order achieved by the MB-MMSE-DF detector. Simulation results show that the MB-MMSE-DF detector achieves a performance superior to existing suboptimal detectors and close to the MLD, while requiring significantly lower complexity.

Index Terms—MIMO systems, spatial multiplexing, decision feedback receivers, iterative methods.

I. INTRODUCTION

THE deployment of multiple transmit and receive antennas in wireless communication systems can offer significant multiplexing [1], [2] and diversity gains [3], [4]. The multiplexing gains enable high spectral efficiencies, whereas the diversity gains increase the reliability of the links and provide low error rates. In multi-input multi-output (MIMO) systems, the transmitter and the receiver should be appropriately designed in order to exploit the structure of the propagation channels. In a spatial multiplexing configuration, the capacity gain grows linearly with the minimum number of transmit and receive antennas [1], [2]. In this scenario, the system can obtain substantial gains in data rate with the transmission of individual data streams from the transmitter to the receiver.

In order to separate these streams, a designer must resort to MIMO detection techniques, which are similar to multiuser detection methods [5]. The optimal maximum likelihood (ML) detector is too complex to be implemented in systems with a large number of antennas. The ML solution can be alternatively computed using sphere decoder (SD) algorithms [6]-[12], which are very efficient for MIMO systems with a small number of antennas. However, the computational complexity of SD algorithms depends on the noise variance, the number of data streams to be detected and the signal constellation, resulting in high computational costs for low signal-to-noise ratios (SNR), large MIMO systems and high-order constellations. The high computational complexity of the ML detector and the SD algorithms in some of the aforementioned situations have motivated the development of numerous alternative strategies for MIMO detection. The linear detector [13], the successive interference cancellation (SIC) approach used in the Vertical-Bell Laboratories Layered Space-Time (VBLAST) systems [14]-[16] and other decision-driven detectors such as decision feedback (DF) [17]-[49] are techniques that can offer attractive trade-offs between performance and complexity. Prior work on DF schemes has been reported with DF detectors with SIC (S-DF) [17], [49], [28] and DF receivers with parallel interference cancellation (PIC) (P-DF) [31], [32], combinations of these schemes [31], [34], [35], [37] and mechanisms to mitigate error propagation [38], [39]. An often criticized aspect of these sub-optimal schemes is that they typically do not achieve the full receive-diversity order of the ML algorithm. This has motivated the investigation of alternative detection strategies such as lattice-reduction (LR) schemes [22]-[23], QR decomposition and the M-algorithm (QRD-M) detectors [24], [25] , probabilistic data association (PDA) [26], [27] detectors, extensions to soft-input soft-output detectors [46]-[50], and calls for flexible cost-effective detection algorithms with near-ML or ML performance, which achieve the full receive-diversity order.

In this work, a DF detection strategy based on multiple branches (MB) is proposed for MIMO systems operating in a spatial multiplexing configuration. The proposed detection algorithm, termed as MB-MMSE-DF and first reported in [36], employs multiple feedforward and feedback receive filters with appropriate transformations that are obtained from a common matrix inverse and allow the search for improved detection candidates. To this end, the MB-MMSE-DF receiver exploits different patterns and orderings, and selects the branch

with the highest likelihood based on an instantaneous MMSE metric. Constrained minimum mean-squared error (MMSE) receive filters designed with constraints on the shape and magnitude of the feedback filters for the proposed MB-MMSE-DF receiver are devised. The MB-MMSE-DF detector does not require a significant additional complexity over the conventional MMSE-DF receiver since it relies on filter realizations with different constraints on the feedback filters, a common matrix inversion and the same second-order statistics. An adaptive implementation of the MB-MMSE-DF detector with a recursive least squares (RLS)-type algorithm for estimating the parameters of the filters when the channel is time-varying is also presented. The optimal ordering algorithm for the MB-MMSE-DF detector is presented along with a low-complexity suboptimal ordering technique. A soft-input soft-output version of the MB-MMSE-DF receiver for iterative detection and decoding using convolutional codes is also developed. The iterative MB-MMSE-DF receiver employs multiple detection candidates to construct a list of log-likelihood ratios for each transmitted bit. A diversity analysis that discusses the diversity order achieved by the MB-MMSE-DF detector is carried out along with a computational complexity study. The MB-MMSE-DF detector achieves a performance close to the optimal ML detector, while it requires a reduced cost and has a superior performance to existing sub-optimal detectors.

The main contributions of this work are:
1) The proposal of the MB-MMSE-DF detection algorithm;
2) MMSE expressions for filter design along with shape patterns and magnitude constraints for the filters;
3) An adaptive version of the proposed detection scheme along with a performance and a complexity analysis;
4) An optimal ordering algorithm is presented along with a cost-effective suboptimal ordering algorithm for the MB-MMSE-DF detector;
5) An iterative MB-MMSE-DF algorithm for processing soft estimates with convolutional codes;
6) An analysis of the complexity and diversity order attained by the MB-MMSE-DF detector;
7) A comparative study of the MB-MMSE-DF and existing MIMO detection algorithms.

This paper is organized as follows. Section II briefly describes a MIMO spatial multiplexing system model. Section III is devoted to the proposed MB-MMSE-DF detection algorithm, the design of the MMSE filters and a multistage scheme. Section IV presents the design of the shaping matrices, the ordering and the parameter estimation algorithms. Section V is dedicated to the development of an iterative version of the MB-MMSE-DF detector which processes soft information for iterative detection and decoding. Section VI presents an analysis of the computational complexity along with the diversity order of the MB-MMSE-DF scheme. Section VII presents and discusses the simulation results and Section VIII draws the conclusions.

II. SYSTEM MODEL

Consider a spatial multiplexing MIMO system with N_T transmit antennas and N_R receive antennas, where $N_R \geq N_T$. At each time instant i, the system transmits N_T symbols which are organized into a $N_T \times 1$ vector $\boldsymbol{s}[i] =$
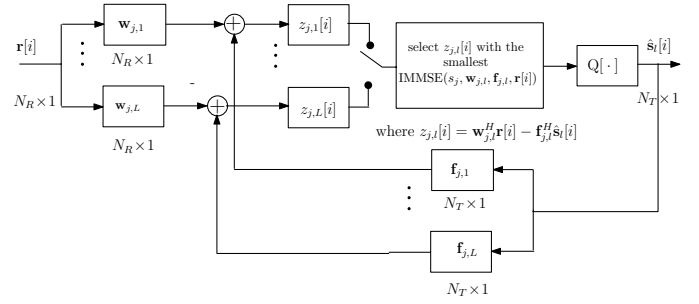

Fig. 1. Block diagram of the proposed MB-MMSE-DF detector and the processing of the jth data stream.

$\begin{bmatrix} s_1[i], & s_2[i], & \ldots, & s_{N_T}[i] \end{bmatrix}^T$ taken from a modulation constellation $A = \{a_1, a_2, \ldots, a_N\}$, where $(\cdot)^T$ denotes transpose and $N = 2^C$. In other words, each symbol is carrying C bits. The symbol vector $\boldsymbol{s}[i]$ is then transmitted over flat fading channels and the signals are demodulated and sampled at the receiver, which is equipped with N_R antennas.

The received signal after demodulation, matched filtering and sampling is organized in an $N_R \times 1$ vector $\boldsymbol{r}[i] = \begin{bmatrix} r_1[i], & r_2[i], & \ldots, & r_{N_R}[i] \end{bmatrix}^T$ with sufficient statistics for detection as given by

$$\boldsymbol{r}[i] = \boldsymbol{H}\boldsymbol{s}[i] + \boldsymbol{n}[i], \tag{1}$$

where the $N_R \times 1$ vector $\boldsymbol{n}[i]$ is a zero mean complex circular symmetric Gaussian noise with covariance matrix $E\begin{bmatrix} \boldsymbol{n}[i]\boldsymbol{n}^H[i] \end{bmatrix} = \sigma_n^2 \boldsymbol{I}$, where $E[\cdot]$ stands for expected value, $(\cdot)^H$ denotes the Hermitian operator, σ_n^2 is the noise variance and $\boldsymbol{I}$ is the identity matrix. The symbol vector $\boldsymbol{s}[i]$ has zero mean and a covariance matrix $E\begin{bmatrix} \boldsymbol{s}[i]\boldsymbol{s}^H[i] \end{bmatrix} = \sigma_s^2 \boldsymbol{I}$, where σ_s^2 is the signal power. The elements h_{n_R,n_T} of the $N_R \times N_T$ channel matrix $\boldsymbol{H}$ correspond to the complex channel gains from the n_Tth transmit antenna to the n_Rth receive antenna.

III. MULTI-BRANCH MMSE DECISION FEEDBACK DETECTION

In this section, the structure of the proposed MB-MMSE-DF detector for MIMO systems is presented and a schematic of the detector is shown in Fig. 1. The MB-MMSE-DF detector employs multiple pairs of MMSE receive filters in such a way that the detector can obtain different local maxima of the likelihood function and select the best candidate for detection according to an instantaneous MMSE metric for each received data symbol. The receive filters are designed based on the MMSE statistical criterion whereas the detection and the selection of the best candidate for each received symbol relies on an instantaneous MMSE criterion. The MB-MMSE-DF scheme is flexible and approaches the full receive diversity available in the system by increasing the number of branches. The MB-MMSE-DF detector employs tasks such as MB processing, MMSE decision feedback, and ordering that have a combined computational cost that is substantially lower than the ML detector, which is very simple from a mathematical point of view but requires a number of operations that is much higher than the MB-MMSE-DF and other existing detectors.

In order to detect each transmitted data stream using the proposed MB-MMSE-DF detector, the receiver linearly

combines the feedforward filter represented by the $N_R \times 1$ vector $\boldsymbol{w}_{j,l}$ corresponding to the j-th data stream and the l-th branch with the received vector $\boldsymbol{r}[i]$, subtracts the remaining interference by linearly combining the feedback filter denoted by the $N_T \times 1$ vector $\boldsymbol{f}_{j,l}$ with the $N_T \times 1$ vector of initial decisions $\hat{\boldsymbol{s}}_l[i]$ obtained from previous decisions. This process is repeated for L candidate symbols and N_T data streams as described by

$$z_{j,l}[i] = \boldsymbol{w}_{j,l}^H \boldsymbol{r}[i] - \boldsymbol{f}_{j,l}^H \hat{\boldsymbol{s}}_l[i],$$
$$j = 1, \ldots, N_T \ \text{and} \ l = 1, \ldots, L \quad (2)$$

where the input to the decision device for the ith symbol and the j-th stream is the $L \times 1$ vector $\boldsymbol{z}_j[i] = [z_{j,1}[i] \ \ldots \ z_{j,L}[i]]^T$. The number of parallel branches L that produce detection candidates is a parameter that must be chosen by the designer and is determined experimentally. Another important design aspect that affects the performance is the ordering algorithm which will be discussed later on. The goal of this work is to employ a reduced number of branches and yet achieve near-ML or ML performance.

The MB-MMSE-DF detector generates L candidate symbols for each data stream and then selects the best branch according to an instantaneous MMSE metric as described by

$$l_{j,\text{opt}} = \arg \min_{1 \le l_j \le L} \text{IMMSE}(s_j[i], \boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{r}[i]),$$
$$j = 1, \ldots, N_T \quad (3)$$

where

$$\text{IMMSE}(s_j[i], \boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{r}[i])$$
$$\approx |s_j[i]|^2 - \boldsymbol{w}_{j,l}^H \hat{\boldsymbol{R}}[i] \boldsymbol{w}_{j,l} + \boldsymbol{f}_{j,l}^H \hat{\boldsymbol{s}}_l[i] \hat{\boldsymbol{s}}_l^H[i] \boldsymbol{f}_{j,l} \quad (4)$$

where the instantaneous MMSE metric IMMSE is produced by the pair of receive filters $\boldsymbol{w}_{j,l}$ and $\boldsymbol{f}_{j,l}$, the quantity $|s_j[i]|^2$, the received vector $\boldsymbol{r}[i]$ and an instantaneous estimate of the covariance matrix $\hat{\boldsymbol{R}}[i] = \boldsymbol{r}[i] \boldsymbol{r}^H[i]$. Further details about the MMSE and IMMSE expressions are included in the Appendices.

The final detected symbol of the MB-MMSE-DF detector is obtained by using the best branch as given by

$$\hat{s}_j[i] = Q\big[\boldsymbol{z}_{j,l_{j,\text{opt}}}[i]\big]$$
$$= Q\big[\boldsymbol{w}_{j,l_{j,\text{opt}}}^H \boldsymbol{r}[i] - \boldsymbol{f}_{j,l_{j,\text{opt}}}^H \hat{\boldsymbol{s}}_{l_{j,\text{opt}}}[i]\big],$$
$$j = 1, \ldots, N_T \quad (5)$$

where $Q(\cdot)$ is a slicing function that makes the decisions about the symbols, which can be drawn from an M-PSK or a QAM constellation.

A. MMSE Filter Design

In this part, the design of the MMSE receive filters of the proposed MB-MMSE-DF detector is detailed by first assuming imperfect feedback of the symbol decisions ($\boldsymbol{s} \ne \hat{\boldsymbol{s}}$) and then by assuming perfect feedback ($\boldsymbol{s} = \hat{\boldsymbol{s}}$). The design of the receive filters is equivalent to determining feedforward filters $\boldsymbol{w}_{j,l}$ with N_R coefficients and feedback filters $\boldsymbol{f}_{j,l}$ with

N_T elements subject to certain shape constraints on $\boldsymbol{f}_{j,l}$ in accordance to the following optimization problem

$$\min \text{MSE}(s_j[i], \boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l})$$
$$= E\big[|s_j[i] - \boldsymbol{w}_{j,l}^H \boldsymbol{r}[i] + \boldsymbol{f}_{j,l}^H \hat{\boldsymbol{s}}_l[i]|^2\big]$$
$$\text{subject to } \boldsymbol{S}_{j,l} \boldsymbol{f}_{j,l} = \boldsymbol{0} \ \text{ and } \ ||\boldsymbol{f}_{j,l}||^2 = \gamma_{j,l} ||\boldsymbol{f}_{j,l}^c||^2,$$
$$\text{for } j = 1, \ldots, N_T \text{ and } l = 1, \ldots, L,$$
$$(6)$$

where the $N_T \times N_T$ shape constraint matrix is $\boldsymbol{S}_{j,l}$, $\boldsymbol{0}$ is a $N_T \times 1$ constraint vector and $\gamma_{j,l}$ is a design parameter that ranges from 0 to 1 and is responsible for scaling the norm of the conventional feedback receive filter $\boldsymbol{f}_{j,l}^c$. The scaling of $\boldsymbol{f}_{j,l}^c$ results in the desired feedback receive filter $\boldsymbol{f}_{j,l}$. The expectation operator is taken over the random parameters $\boldsymbol{s}[i]$ and $\boldsymbol{r}[i]$ assuming that $\boldsymbol{n}[i]$ and $\boldsymbol{s}[i]$ are statistically independent, and that the entries of $\boldsymbol{s}[i]$ and $\boldsymbol{n}[i]$ are independent and identically distributed random variables. The role of the shape constraint matrix $\boldsymbol{S}_{j,l}$ is to choose the feedback connections which will be used in the interference cancellation. If a designer employs multiple branches and shape constraint matrices along with different orderings then multiple candidates for detection can be generated, resulting in an improved receiver performance. The rationale for scaling the norm of the feedback filter is to reduce the impact of the error propagation and improve the performance of the receiver. This is accomplished by judiciously adjusting the scaling of the norm and employing the value which minimizes the error propagation.

In what follows, the optimal MMSE receive filters based on the proposed optimization in (6) are derived. By resorting to the method of Lagrange multipliers, computing the gradient vectors of the Lagrangian with respect to $\boldsymbol{w}_{j,l}$ and $\boldsymbol{f}_{j,l}$, equating them to null vectors and rearranging the terms, we obtain for $j = 1, \ldots, N_T$ and $l = 1, \ldots, L$

$$\boldsymbol{w}_{j,l}^{\text{MMSE}} = \boldsymbol{R}^{-1}(\boldsymbol{p}_j + \boldsymbol{Q}\boldsymbol{f}_{j,l}), \quad (7)$$

$$\boldsymbol{f}_{j,l}^{\text{MMSE}} = \frac{\beta_{j,l}}{\sigma_s^2} \boldsymbol{\Pi}_{j,l}(\boldsymbol{Q}^H \boldsymbol{w}_{j,l} - \boldsymbol{t}_j), \quad (8)$$

where

$$\boldsymbol{\Pi}_{j,l} = \boldsymbol{I} - \boldsymbol{S}_{j,l}^H (\boldsymbol{S}_{j,l}^H \boldsymbol{S}_{j,l})^{-1} \boldsymbol{S}_{j,l} \quad (9)$$

is a projection matrix that ensures the shape constraint $\boldsymbol{S}_{j,l}$ on the feedback filter, $\beta_{j,l} = (1 - \mu_{j,l})^{-1}$ is the parameter that controls the ability of the MB-MMSE-DF detector to mitigate error propagation with values $0 \le \beta_{j,l} \le 1$, and $\mu_{j,l}$ is the Lagrange multiplier. It should be remarked that the inverse $(\boldsymbol{S}_{j,l}^H \boldsymbol{S}_{j,l})^{-1}$ might not exist. In these situations, a pseudo-inverse is computed. The relationship between $\beta_{j,l}$ and $\gamma_{j,l}$ is not in closed-form except for the extreme values when we have $\beta_{j,l} = 0$ and $\beta_{j,l} = 1$ for $\gamma_{j,l} = 0$ (standard linear MMSE detector) and $\gamma_{j,l} = 1$ (standard MB-MMSE-DF detector), respectively. The optimization of the parameter $\beta_{j,l}$ has been done with the aid of simulations because there is no closed-form solution to obtain $\beta_{j,l}$. The simulation approach has indicated that the performance is improved for a range of parameters between 0.6 and 0.7. This range of parameters was verified to consistently produce good results for all the

scenarios investigated with the MB-MMSE-DF detector. The $N_R \times N_R$ covariance matrix of the input data vector is $\boldsymbol{R} = E[\boldsymbol{r}[i]\boldsymbol{r}^H[i]]$, $\boldsymbol{p}_j = E[\boldsymbol{r}[i]s_j^*[i]]$, $\boldsymbol{Q} = E[\boldsymbol{r}[i]\hat{\boldsymbol{s}}_l^H[i]]$, and $\boldsymbol{t}_j = E[\hat{\boldsymbol{s}}_l[i]s_j^*[i]]$ is the $N_T \times 1$ vector of correlations between $\hat{\boldsymbol{s}}_l[i]$ and $s_j^*[i]$. Substituting (8) into (7) and then further manipulating the expressions we arrive at the following MMSE receive filter expressions

$$\boldsymbol{w}_{j,l}^{\mathrm{MMSE}} = \left(\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H\right)^{-1}\left(\boldsymbol{p}_j - \beta_{j,l}\boldsymbol{\Pi}_{j,l}\boldsymbol{t}_j\right), \quad (10)$$

$$\begin{aligned}
\boldsymbol{f}_{j,l}^{\mathrm{MMSE}} = \frac{\beta_{j,l}}{\sigma_s^2}\boldsymbol{\Pi}_{j,l}&\big(\boldsymbol{Q}^H\left(\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H\right)^{-1} \\
&\left(\boldsymbol{p}_j - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{t}_j\right) - \boldsymbol{t}_j\big).
\end{aligned} \quad (11)$$

The above expressions only depend on statistical quantities, and consequently on the channel matrix $\boldsymbol{H}$, the symbol and noise variances σ_s^2 and σ_n^2, respectively, and the constraints. However, the matrix inversion required for computing $\boldsymbol{w}_{j,l}$ is different for each branch and data stream, thereby rendering the scheme computationally less efficient. The expressions obtained in (7) and (8) are equivalent to those in (10) and (11), and only require iterations between them for an equivalent performance. A key advantage of using (7) and (8) is that they only require a single matrix inversion that is common to all branches and two iterations prior to their use, whereas in (10) and (11) there is a matrix inversion associated with each branch. For this reason, in what follows the expressions in (7) and (8) are adopted and further simplified.

As briefly explained above, the expressions in (7) and (8) can be simplified by evaluating the expected values. By using the fact that $\boldsymbol{t}_j = \boldsymbol{0}$ for interference cancellation as $\hat{\boldsymbol{s}}_l[i]$ does not contain s_j and assuming perfect feedback ($\boldsymbol{s} = \hat{\boldsymbol{s}}$), the following expressions are obtained

$$\boldsymbol{w}_{j,l}^{\mathrm{MMSE}} = \left(\boldsymbol{H}\boldsymbol{H}^H + \sigma_n^2/\sigma_s^2\boldsymbol{I}\right)^{-1}\boldsymbol{H}(\boldsymbol{\delta}_j + \boldsymbol{f}_{j,l}), \quad (12)$$

$$\boldsymbol{f}_{j,l}^{\mathrm{MMSE}} = \beta_{j,l}\boldsymbol{\Pi}_{j,l}\boldsymbol{H}^H\boldsymbol{w}_{j,l}, \quad (13)$$

where $\boldsymbol{\delta}_j = [\underbrace{0\ldots0}_{j-1}\ 1\ \underbrace{0\ldots0}_{N_T-j-2}]^T$ is a $N_T \times 1$ vector with a one in the jth element and zeros elsewhere. A step-by-step derivation of the filters is shown in Appendix I. The proposed MB-MMSE-DF detector expressions above require the channel matrix $\boldsymbol{H}$ (in practice an estimate of it) and the noise variance σ_n^2 at the receiver. In terms of complexity, it requires for each branch l the inversion of an $N_R \times N_R$ matrix and other operations with complexity $O(N_R^3)$. However, the expressions obtained in (7) and (8) for the general case, and in (12) and (13) for the case of perfect feedback, reveal that the most expensive operations, i.e., the matrix inversions, are identical for all branches. Therefore, the design of receive filters for the multiple branches only requires further additions and multiplications of the matrices. Moreover, it can be verified that the filters $\boldsymbol{w}_{j,l}^{\mathrm{MMSE}}$ and $\boldsymbol{f}_{j,l}^{\mathrm{MMSE}}$ are dependent on one another, which means the designer has to iterate them before applying the detector. It has been verified by simulations with different system parameters and by comparing the resulting parameters of the receive filters with those obtained by (10) and (11) that it suffices to employ two iterations of (12) and (13) to have a performance equivalent to that obtained by using (10) and (11). For this reason, we employ the receive filters of
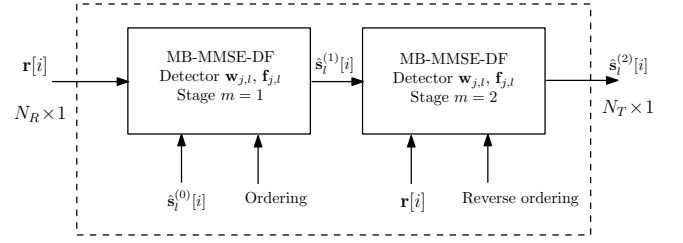

Fig. 2. Block diagram of the proposed two-stage MB-MMSE-DF detector.

(12) and (13) with two iterations in the proposed MB-MMSE-DF detector.

The MMSE associated with the filters $\boldsymbol{w}_{j,l}^{\mathrm{MMSE}}$ and $\boldsymbol{f}_{j,l}^{\mathrm{MMSE}}$ and the statistics of the data symbols $s_j[i]$ is given by

$$\underbrace{\mathrm{MMSE}(s_j[i], \boldsymbol{w}_{j,l}^{\mathrm{MMSE}}, \boldsymbol{f}_{j,l}^{\mathrm{MMSE}})}_{\mathrm{MMSE}_j}$$
$$= \sigma_s^2 - \boldsymbol{w}_{j,l}^{H,\ \mathrm{MMSE}}\boldsymbol{R}\boldsymbol{w}_{j,l}^{\mathrm{MMSE}} + \boldsymbol{f}_{j,l}^{H,\ \mathrm{MMSE}}\boldsymbol{f}_{j,l}^{\mathrm{MMSE}}, \quad (14)$$

where $\sigma_s^2 = E[|s_j[i]|^2]$ is the variance of the desired symbol. A detailed derivation of the MMSE associated with the receive filters is shown in Appendix II along with connections with the MMSE achieved by conventional DF detectors.

B. Multi-stage Detection for the MB-MMSE-DF

In this subsection, algorithms for error propagation mitigation are presented and incorporated into the structure of the MB-MMSE-DF detection scheme. The strategy is based on iterative multi-stage detection [31], [34] that gradually refines the decision vector and improves the overall performance. It is incorporated into the MB-MMSE-DF scheme and the improvements in the detection performance are then investigated.

The basic principle underlying multi-stage detection is to iteratively refine the estimates of the decision vector used in DF receivers [31], [34] and mitigate error propagation. An advantage of multi-stage detection that has not been exploited for the design of MIMO detectors is that of equalizing the performance of the detectors over the data streams. Since V-BLAST or DF detection usually favors certain data streams (the last detected ones) with respect to performance, this might be important for some applications where fairness or uniform performance is required between the data streams. This concept is incorporated into the proposed MB-MMSE-DF scheme and the MMSE design of the previous subsection. An MB-MMSE-DF scheme is employed in each stage and the estimates of the decision vector are gradually refined as illustrated in Fig. 2. Specifically, a multi-stage algorithm for the MB-MMSE-DF can be described by

$$z_{j,l}^{(m+1)}(i) = \boldsymbol{w}_{j,l}^{H,\ \mathrm{MMSE}}\boldsymbol{r}[i] - \boldsymbol{f}_{j,l}^{H,\ \mathrm{MMSE}}\hat{\boldsymbol{s}}_l^{(m)}[i],$$
$$m = 0,\ 1,\ \ldots,\ M, \quad (15)$$

where the MMSE filters $\boldsymbol{w}_{j,l}^{\mathrm{MMSE}}$ and $\boldsymbol{f}_{j,l}^{\mathrm{MMSE}}$ are designed with the approach detailed in the previous subsection, M denotes the number of stages and $\hat{\boldsymbol{s}}_l^{(m)}[i]$ is the vector of tentative decisions from the preceding iteration that is described by

$$\hat{s}_{j,l}^{(0)}[i] = \mathrm{Q}\left(\boldsymbol{w}_{j,l}^{H,\ \mathrm{MMSE}}\boldsymbol{r}[i]\right), \quad j, l = 1,\ \ldots,\ N_T, \quad (16)$$

$$\hat{s}_{j,l_{j,\mathrm{opt}}}^{(m)}[i] = \mathrm{Q}\left(z_{j,l_{j,\mathrm{opt}}}^{(m)}[i]\right), \quad m = 1, \ \ldots, \ M, \quad (17)$$

where the number of stages M depends on the scenario.

In order to equalize the performance over the data streams, an M-stage structure is considered. The first stage is an MB-MMSE-DF scheme with filters $\boldsymbol{w}_{j,l}^{\mathrm{MMSE}}$ and $\boldsymbol{f}_{j,l}^{\mathrm{MMSE}}$. The tentative decisions are passed to the second stage, which consists of another MB-MMSE-DF scheme with the same receive filters that uses the decisions of the first stage and so successively. The resulting multi-stage MIMO detection scheme is denoted I-MB-MMSE-DF. The output of the second stage of the resulting scheme is

$$z_{j,l}^{(2)}[i] = [\boldsymbol{T}\boldsymbol{w}_{j,l}^{\mathrm{MMSE}}]^{H}\boldsymbol{r}[i] - [\boldsymbol{T}\boldsymbol{f}_{j,l}^{\mathrm{MMSE}}]^{H}\hat{\boldsymbol{s}}_{l_{j,\mathrm{opt}}}^{(1)}[i], \quad (18)$$

where $z_{j,l}^{(2)}[i]$ is the output of jth data stream after multi-stage detection with $M = 2$ stages, $\boldsymbol{T}$ is a square permutation matrix with ones along the reverse diagonal and zeros elsewhere. When multiple stages are used, it is beneficial to demodulate the data streams successively and in reverse order relative to the first branch of the MB-MMSE-DF detector. The role of reversing the cancellation order in successive stages is to equalize the performance of the users over the population or at least reduce the performance disparities. It provides a better performance than keeping the same ordering as the last decoded users in the first stage tend to be favored by the reduced interference. The rationale is that the performance can be improved by using the data streams that benefited from interference cancellation (last decoded ones) as the first ones to be decoded in the second stage. Additional stages can be included, although the results suggest that the gains in performance are marginal for more than two stages. Hence, the two-stage scheme is adopted for the rest of this work.

IV. DESIGN OF CANCELLATION PATTERNS, ORDERING AND ADAPTIVE ALGORITHMS

In this section, the design of the shape constraint matrices $\boldsymbol{S}_{j,l}$ is detailed and their choices are motivated. An optimal and a suboptimal ordering algorithms are described for the interference cancellation. An adaptive version of the MB-MMSE-DF detector with RLS-type algorithms is also devised.

A. Design of Cancellation Patterns

The idea of the shape constraint matrices $\boldsymbol{S}_{j,l}$ is to modify the structure of the feedback filters $\boldsymbol{f}_{j,l}$ in such a way that only the selected feedback elements of $\boldsymbol{f}_{j,l}$ will be used to cancel the interference between the data streams. The feedback connections perform interference cancellation with a chosen ordering. If a designer employs multiple branches and shape constraint matrices along with different orderings then multiple candidates for detection can be generated, resulting in an improved receiver performance. The matrices $\boldsymbol{S}_{j,l}$ for the N_T data streams and for the L branches of the MB-MMSE-DF detector can be stored at the receiver and used either online or offline in the design of the feedback filters $\boldsymbol{f}_{j,l}$. In particular, with this approach the ML solution can be searched from different points of the likelihood function using an MMSE-type detector as the starting point. Specifically, the aim is to design and shape the filters $\boldsymbol{f}_{j,l}$ for the N_T data streams and

the L branches with the $N_T \times N_T$ matrices $\boldsymbol{S}_{j,l}$ such that constraint vector is a null vector. This corresponds to allowing feedback connections of only a subgroup of data streams. For the first branch of detection ($l = 1$), the successive cancellation used in the VBLAST [14] can be employed which corresponds mathematically to

$$\boldsymbol{S}_{j,l}\boldsymbol{f}_{j,l} = \boldsymbol{0}, \quad l = 1$$
$$\boldsymbol{S}_{j,l} = \begin{bmatrix} \boldsymbol{0}_{N_T-j+1,N_T-j+1} & \boldsymbol{0}_{N_T-j+1,j-1} \\ \boldsymbol{0}_{j-1,N_T-j+1} & \boldsymbol{I}_{j-1,j-1} \end{bmatrix}, \quad (19)$$
$$j = 1, \ldots, N_T,$$

where $\boldsymbol{0}_{m,n}$ denotes an $m \times n$-dimensional matrix full of zeros, and $\boldsymbol{I}_m$ denotes an m-dimensional identity matrix. Interestingly, when detecting a data stream of interest the feedback connection associated with it cannot be used to subtract interference because it will simply cancel the data stream of interest itself. This is well known in the literature of decision feedback receivers [31], [34] and is the reason for using these structures with constraints. For the remaining branches, an approach based on permutations of the structure of the matrices $\boldsymbol{S}_{j,l}$ is adopted, which is given by

$$\boldsymbol{S}_{j,l}\boldsymbol{f}_{j,l} = \boldsymbol{0}, \quad l = 2, \ldots, L$$
$$\boldsymbol{S}_{j,l} = \phi_l \begin{bmatrix} \boldsymbol{0}_{N_T-j+1,N_T-j+1} & \boldsymbol{0}_{N_T-j+1,j-1} \\ \boldsymbol{0}_{j-1,N_T-j+1} & \boldsymbol{I}_{j-1,j-1} \end{bmatrix}, \quad (20)$$
$$j = 1, \ldots, N_T,$$

where the operator $\phi_l[\cdot]$ permutes the elements of the argument matrix such that this results in different cancellation patterns. For instance, the non-zero elements of the feedback filter $\boldsymbol{f}_{j,l}$ are chosen according to the shape constraint matrices. The permutations for the different branches will change the non-zero elements and allow the receiver to obtain detection candidates from different interference cancellation patterns. Although the above structure is imposed to determine the number of feedback connections for each data stream, it might result in a projection matrix $\boldsymbol{\Pi}_{j,l}$ whose inverse $(\boldsymbol{S}_{j,l}^{H}\boldsymbol{S}_{j,l})^{-1}$ does not exist. In these situations, a pseudo-inverse is computed. While there are different permutations or rotations, the permutations employed are straightforward to implement and will simply change the positions of the non-zero coefficients of the feedback filters. Specifically, the permutation implemented by the function $\phi_l[\cdot]$ is employed together with the ordering to generate a list of candidates for detection. The MB-MMSE detector then chooses a candidate out of L branches for each data stream which benefits from the interference cancellation, thereby processing a data stream that is free or has a reduced level of interference. This increases the diversity order of the MB-MMSE-DF detector, as will be explained in the analysis of the MB-MMSE-DF detector.

An alternative approach for shaping the constraint matrices $\boldsymbol{S}_{j,l}$ for one of the L branches is to use a parallel interference cancellation (PIC) approach [31] and design the matrices as follows

$$\boldsymbol{S}_{j,l}\boldsymbol{f}_{j,l} = \boldsymbol{0}, \quad l = 1, \ 2, \ \ldots, \ L,$$
$$\boldsymbol{S}_{j,l} = \boldsymbol{I}_{N_T} - \mathrm{diag}\,(\boldsymbol{\delta}_j), \quad j = 1, \ldots, N_T, \quad (21)$$

where $\boldsymbol{\delta}_j$ is an $N_T \times 1$ vector with a one in the j-th position and zeros elsewhere. The PIC requires an initial vector of decisions

obtained with the feedforward filters $\boldsymbol{w}_{j,l}$. A problem with the PIC approach of [31] is that it is prone to error propagation due to the cancellation of all but the stream of interest.

B. Ordering Algorithms

The aim of an ordering algorithm in a MIMO system is to obtain a sequence for interference cancellation that optimizes a given criterion. For a conventional SIC detector, the optimal ordering algorithm must test $N_T!$ possibilities with the objective of minimizing the BER [28], [53]. A common alternative to this exhaustive search is to employ a technique based on the norm of the channels, the MMSE or the signal-to-interference-plus-noise ratio (SINR) [1], [14]. The proposed MB-MMSE-DF detector operates with an ordering based on the MMSE and the goal is to find the best performing set of orderings over L branches. The optimal ordering algorithm for the MB-MMSE-DF detector with $L > 1$, which minimizes the MMSE for each data stream, requires testing $\underbrace{N_T!.N_T!\ldots N_T!}_{L}$

possibilities and is given by

$$\{o_{1,l}, \ldots, o_{N_T,l}\}_{\text{opt}} = \arg \min_{o_{1,l},\ldots,o_{N_T,l}} \sum_{l=1}^{L} \sum_{j=1}^{N_T} \text{MMSE}_j,$$
$$\text{for } l = 1, \ldots, L.$$
(22)

The rationale for this algorithm is to find the optimal ordering for each branch, which employs the MMSE over the L branches to find the best performing set of orderings. Again, this requires testing $\underbrace{N_T!.N_T!\ldots N_T!}_{L}$ possibilities. The computational complexity of the algorithm in (22) can increase significantly for large N_T and L. For this reason, a suboptimal ordering algorithm is also presented for the MB-MMSE-DF detector.

In the proposed suboptimal ordering algorithm, a simplified strategy is presented based on the maximization of the difference between the MMSE values obtained for each data stream. For the first branch, an ordering algorithm based on increasing values of the MMSE is considered, and this is equivalent to an ordering according to the maximization of the SINR for a single branch. The ordering of the remaining branches (for the case with $L > 1$) depends on the maximization of the difference between the MMSE of different data streams and is given by

$$o_{j,l} = \arg \max_n \sum_{q=1}^{j-1} |\text{MMSE}_n - \text{MMSE}_{o_{j,q}}|,$$
$$\text{for } l = 2, \ldots, L \text{ and } j, n = 1, \ldots, N_T$$
$$\text{subject to } \text{MMSE}_{o_{j,l}} \neq \text{MMSE}_{o_{q,l}}, \; q = 1, \ldots, j-1.$$
(23)

The principle behind the ordering given by (23) with the multiple branches is to benefit a given data stream or group for each decoding branch. Following this approach, a data stream that for a given ordering appears to be in an unfavorable scenario (with more interference) can benefit in other parallel branches by being detected in a situation with

less interference, increasing the diversity order of the MB-MMSE-DF detector. In other words, the algorithm attempts to obtain orderings that are associated with the largest Euclidean distance between values of MMSE for each data stream as illustrated in Table I. This heuristic turns out to work very well as it will be shown later. The ordering algorithm in (23) requires a number of operations (subtractions, modulus, and comparisons) that are linear in the number of data streams (N_T) and branches (L), i.e., $O(N_T L)$. In the case of static channels, the ordering algorithms can be employed only once at the beginning of the transmission. In the case of time-varying channels and whenever the MMSE obtained changes, the ordering algorithms need to be re-computed in order to ensure an optimized performance.

C. Adaptive MB-MMSE-DF with RLS Algorithms

In this part, an adaptive version of the MB-MMSE-DF detector with an RLS-type algorithm is developed. The aim is to reduce the required computational complexity of the expressions in (7) and (8) from $O(N_R^3)$ to $O(N_R^2)$, and equip the proposed MB-MMSE-DF detector with the ability to track time-varying channels. The procedure to estimate $\boldsymbol{R}^{-1}$ employs the matrix inversion lemma [40]:

$$\boldsymbol{k}[i] = \frac{\lambda^{-1}\hat{\boldsymbol{R}}[i-1]\boldsymbol{r}[i]}{1 + \lambda^{-1}\boldsymbol{r}^H[i]\hat{\boldsymbol{R}}[i-1]\boldsymbol{r}[i]}, \quad (24)$$

$$\hat{\boldsymbol{R}}[i] = \lambda^{-1}\hat{\boldsymbol{R}}[i-1] - \lambda^{-1}\boldsymbol{k}[i]\boldsymbol{r}^H[i]\boldsymbol{R}^{-1}[i-1] \quad (25)$$

where $0 \ll \lambda < 1$ is a forgetting factor that is chosen according to the environment. The estimates of $\boldsymbol{p}_k[i]$ and $\boldsymbol{Q}[i]$ are then computed with the following recursions

$$\hat{\boldsymbol{Q}}[i] = \lambda\hat{\boldsymbol{Q}}[i-1] + \boldsymbol{r}[i]\hat{\boldsymbol{s}}_{l_{\text{opt}}}^H[i], \quad (26)$$

$$\hat{\boldsymbol{p}}_j[i] = \lambda\hat{\boldsymbol{p}}[i-1] + \boldsymbol{r}[i]s_{j,l_{\text{opt}}}^*[i], l = 1, \ldots, L \quad (27)$$

where the decision vector $\hat{\boldsymbol{s}}_{l_{\text{opt}}}[i] = Q(\hat{\boldsymbol{w}}_{j,l}^H[i-1]\boldsymbol{r}[i] - \hat{\boldsymbol{f}}_{j,l}^H[i-1]\hat{\boldsymbol{s}}_l[i])$ is obtained with the filters of the previous time instant. The feedforward filters for $l = 1, \ldots, L$ and $j = 1, \ldots, N_T$ are computed by

$$\hat{\boldsymbol{w}}_{j,l}[i] = \hat{\boldsymbol{R}}^{-1}[i](\hat{\boldsymbol{p}}_j[i] + \hat{\boldsymbol{Q}}[i]\hat{\boldsymbol{f}}_{j,l}[i-1]), \quad (28)$$

Once the feedforward filters are computed the feedback filters can be obtained by

$$\hat{\boldsymbol{f}}_{j,l}[i] = \beta_{j,l}\boldsymbol{\Pi}_{j,l}\hat{\boldsymbol{Q}}^H[i]\hat{\boldsymbol{w}}_{j,l}[i]. \quad (29)$$

Note that the filters need to be initialized and that the computation of $\hat{\boldsymbol{R}}^{-1}[i]$ is common to all branches, i.e., we only need to compute it once for all branches. A summary of the adaptive MB-MMSE-DF detector is given in Table II. The receive filters are computed in an alternating fashion, i.e., one receive filter is updated followed by the other and the cycle is repeated for every data symbol. Note that the RLS-type algorithm presented in Table 1 is a standard version that might need modifications for a numerically-stable hardware implementation. In the case of a hardware implementation, a square root (or QR decomposition) version will have better numerical properties because the square-root structures do not amplify numerical errors and tend to assume values within smaller dynamic ranges [40]. Other advanced algorithms might also be considered [41]-[43].

TABLE I
PROPOSED SUBOPTIMAL ORDERING ALGORITHM.

1. Ordering $O_l = \{o_{1,l} \ldots o_{j,l} \ldots o_{N_T,l}\}$ for branch $l = 1$:
Compute MMSE for each stream: $\text{MMSE}_j = \sigma_j - \boldsymbol{h}_j^H \boldsymbol{R}^{-1} \boldsymbol{h}_j$
Calculate the ordering based on increasing values of MMSE_j:
$o_{j,l} = \arg \min_j \text{MMSE}_j$, for $j = 1, \ldots, N_T$
subject to $\text{MMSE}_j \geq \text{MMSE}_{o_{q,l}}$, $q = 1, \ldots, j-1$
2. Ordering $O_l = \{o_{1,l} \ldots o_{j,l} \ldots o_{N_T,l}\}$ for branches $l = 2, \ldots, L$:
$o_{j,l} = \arg \max_n \sum_{q=1}^{j-1} \lvert \text{MMSE}_n - \text{MMSE}_{o_{j,q}} \rvert$, for $j, n = 1, \ldots, N_T$
subject to $\text{MMSE}_{o_{j,l}} \neq \text{MMSE}_{o_{q,l}}$, $q = 1, \ldots, j-1$

TABLE II
PROPOSED ADAPTIVE MB-MMSE-DF DETECTION ALGORITHM.

1. Initialize parameters: ordering, L, $\boldsymbol{S}_{j,l}$, $\beta_{j,l}$, N_T and N_R.
For $i = 1, \ldots, Q$, where Q is the packet size do
2. Compute $\hat{\boldsymbol{R}}^{-1}[i]$ as follows
$\boldsymbol{k}[i] = \frac{\lambda^{-1} \hat{\boldsymbol{R}}[i-1] \boldsymbol{r}[i]}{1 + \lambda^{-1} \boldsymbol{r}^H[i] \hat{\boldsymbol{R}}[i-1] \boldsymbol{r}[i]}$,
$\hat{\boldsymbol{R}}[i] = \lambda^{-1} \hat{\boldsymbol{R}}[i-1] - \lambda^{-1} \boldsymbol{k}[i] \boldsymbol{r}^H[i] \boldsymbol{R}^{-1}[i-1]$.
3. Obtain $\hat{\boldsymbol{Q}}[i]$ as given by
$\hat{\boldsymbol{Q}}[i] = \lambda \hat{\boldsymbol{Q}}[i-1] + \boldsymbol{r}[i] \hat{\boldsymbol{s}}_{l_{j,\text{opt}}}^H[i]$.
4. Compute $\hat{\boldsymbol{p}}_j[i]$ for $j = 1, \ldots, L$ as follows
$\hat{\boldsymbol{p}}_j[i] = \lambda \hat{\boldsymbol{p}}[i-1] + \boldsymbol{r}[i] s_{j,l_{\text{opt}}}^*[i]$.
5. Determine the ordering $o_{1,l}, \ldots, o_{N_T,l}$ for $l = 1, \ldots, L$
6. For $l = 1, \ldots, L$ and $j = 1, \ldots, N_T$ compute
$\hat{\boldsymbol{w}}_{j,l}[i] = \hat{\boldsymbol{R}}^{-1}[i] \hat{\boldsymbol{H}}[i] (\boldsymbol{\delta}_j + \hat{\boldsymbol{f}}_{j,l}[i-1])$
$\hat{\boldsymbol{f}}_{j,l}[i] = \beta_{j,l} \boldsymbol{\Pi}_{j,l} \hat{\boldsymbol{H}}^H[i] \hat{\boldsymbol{w}}_{j,l}[i]$
7. For $l = 1, \ldots, L$ and $j = 1, \ldots, N_T$ do
Obtain $z_{j,l}[i] = \hat{\boldsymbol{w}}_{j,l}^H[i] \boldsymbol{r}[i] - \hat{\boldsymbol{f}}_{j,l}^H[i] \hat{\boldsymbol{s}}_l[i]$
Determine $l_{j,\text{opt}} = \arg \min_{1 \leq l_j \leq L} \text{IMMSE}(s_j[i], \hat{\boldsymbol{w}}_{j,l}[i], \hat{\boldsymbol{f}}_{j,l}[i])$
Detect symbol: $\hat{s}_j[i] = Q\big[z_{j,l_{\text{opt}}}[i]\big]$

V. ITERATIVE SOFT-INPUT SOFT-OUTPUT DETECTION AND DECODING

This section presents an iterative version of the proposed MB-MMSE-DF detector operating with soft-input soft-output detection and decoding, and with convolutional codes [44]-[49]. The motivation for the proposed scheme is that significant gains can be obtained from iterative techniques with soft cancellation methods and channel codes [44]-[49] when combined with efficient receiver algorithms. A low-complexity iterative MB-MMSE-DF receiver that works with a reduced list of candidate symbols and log-likelihood ratios (LLRs), and that can approach the performance of the optimal detector is developed. The MIMO system described in Section II is considered with convolutional codes. The proposed iterative receiver structure consists of the following stages: a soft-input-soft-output (SISO) MB-MMSE-DF detector and a maximum *a posteriori* (MAP) decoder. The receiver structure also incorporates a selection strategy for the list of LLRs which are used to refine the exchange of soft information. These stages are separated by interleavers and deinterleavers. The soft outputs from the MB-MMSE-DF are used to estimate LLRs which are interleaved and serve as input to the MAP decoder for the convolutional code. The MAP decoder computes *a posteriori* probabilities (APPs) for each stream's encoded symbols, which are used to generate soft estimates. These soft estimates are subsequently used to update the receive filters of the MB-MMSE-DF detector, de-interleaved and fed back through the feedback filter. The MB-MMSE-DF detector computes the *a posteriori* log-likelihood ratio (LLR) of a transmitted symbol (+1 or −1) for every code bit of each data stream as given by

$$\Lambda_1[b_{j,c,l}[i]] = \log \frac{P[b_{j,c,l}[i] = +1 | \mathbf{r}[i]]}{P[b_{j,c,l}[i] = -1 | \mathbf{r}[i]]}, \quad (30)$$
$$j = 1, \ldots, N_T, \ c = 1, \ldots, C, \ l = 1, \ldots, L,$$

where C is the number of bits used to map the constellation. Using Bayes' rule, the above equation can be written as

$$\Lambda_1[b_{j,c,l}[i]] =$$
$$\log \frac{P[\mathbf{r}[i] | b_{j,c,l}[i] = +1]}{P[\mathbf{r}[i] | b_{j,c,l}[i] = -1]} + \log \frac{P[b_{j,c}[i] = +1]}{P[b_{j,c}[i] = -1]} = \quad (31)$$
$$\lambda_1[b_{j,c,l}[i]] + \lambda_2^p[b_{j,c}[i]],$$

where $\lambda_2^p[b_{j,c}[i]] = \log \frac{P[b_{j,c}[i] = +1]}{P[b_{j,c}[i] = -1]}$ is the *a priori* LLR of the code bit $b_{j,c}[i]$, which is computed by the MAP decoder processing the jth data stream in the previous iteration, interleaved and then fed back to the MB-MMSE-DF detector. The superscript ^{p} denotes the quantity obtained in the previous iteration. Assuming equally likely bits, we have $\lambda_2^p[b_{j,c}[i]] = 0$ in the first iteration for all streams. The quantity $\lambda_1[b_{j,c,l}[i]] = \log \frac{P[\mathbf{r}[i] | b_{j,c,l}[i] = +1]}{P[\mathbf{r}[i] | b_{j,c,l}[i] = -1]}$ represents the *extrinsic* information computed by the SISO MB-MMSE-DF detector based on the received data $\mathbf{r}[i]$, and the prior information about the code bits $\lambda_2^p[b_{j,c}[i]]$, $j = 1, \ldots, N_T$, $c = 1, \ldots, C$ and the ith data symbol. Unlike prior work on soft interference cancellation [44], [47], [49] and list sphere decoders [10], [12], [46], the extrinsic information $\lambda_1[b_{j,c,l}[i]]$ is obtained from a list of candidate symbols generated by the MB-MMSE-DF detector and the prior information provided by the MAP decoder, which is de-interleaved and fed back into the MAP decoder of the jth data stream as the *a priori* information in the next iteration.

For the MAP decoding, we assume that the interference plus noise at the output $\boldsymbol{z}_{j,l}[i]$ of the linear receive filters is Gaussian. This assumption has been reported in previous works [44]-[49] and provides an efficient and accurate way of computing the extrinsic information. Thus, for the jth stream, the lth branch and the qth iteration the soft output of the MB-MMSE-DF detector is

$$z_{j,l}^{(q)}[i] = V_{j,l}^{(q)} s_{j,l}[i] + \xi_{j,l}^{(q)}[i], \quad (32)$$

where $V_{j,l}^{(q)}[i]$ is a scalar variable equivalent to the magnitude of the channel corresponding to the jth data stream and $\xi_{j,l}^{(q)}[i]$ is a Gaussian random variable with variance $\sigma_{\xi_{j,l}^{(q)}}^2$. Since we have

$$V_{j,l}^{(q)}[i] = E\big[s_{j,l}^*[i] z_{j,l}^{(q)}[i]\big] \quad (33)$$

and

$$\sigma^2_{\xi^{(q)}_{j,l}}[i] = E\big[|z^{(q)}_{j,l}[i] - V^{(q)}_{j,l}[i]s_{j,l}[i]|^2\big], \qquad (34)$$

the receiver can obtain the estimates $\hat{V}^{(q)}_{j,l}[i]$ and $\hat{\sigma}^2_{\xi^{(q)}_{j,l}}[i]$ via corresponding sample averages over the received symbols. These estimates are used to compute the *a posteriori* probabilities $P[b_{j,c,l}[i] = \pm1|z^{(q)}_{j,l}[i]]$ which are de-interleaved and used as input to the MAP decoder. In what follows, it is assumed that the MAP decoder generates APPs $P[b_{j,c,l}[i] = \pm1]$, which are used to compute the input to the feedback filter $\boldsymbol{f}_{j,l}$. From (32) the extrinsic information generated by the iterative MB-MMSE-DF is given by

$$\lambda_1[b_{j,c,l}[i]] = \log \frac{P[z^{(q)}_{j,l}(i)|b_{j,c,l}[i] = +1]}{P[z^{(q)}_{j,l}[i]|b_{j,c,l}[i] = -1]}$$

$$= \log \frac{\sum\limits_{\mathbb{S} \in \mathbb{S}^{+1}_c} \exp\Big(-\frac{|z^{(q)}_{j,l}[i] - V^{(q)}_{j,l}\mathbb{S}|^2}{2\sigma^2_{\xi^{(q)}_{j,l}}[i]}\Big)}{\sum\limits_{\mathbb{S} \in \mathbb{S}^{-1}_c} \exp\Big(-\frac{|z^{(q)}_{j,l}[i] - V^{(q)}_{j,l}\mathbb{S}|^2}{2\sigma^2_{\xi^{(q)}_{j,l}}[i]}\Big)}, \qquad (35)$$

where $\mathbb{S}^{+1}_c$ and $\mathbb{S}^{-1}_c$ are the sets of all possible constellations that a symbol can take on such that the cth bit is 1 and -1, respectively. Different approaches are possible to compute the extrinsic information generated from the list of soft estimates provided by the iterative MB-MMSE-DF detector. In this work, the iterative MB-MMSE-DF detector chooses the LLR from a list of L candidates for the decoding iteration as

$$\lambda_1[b_{j,c,l_{\text{opt}}}[i]] = \arg \max_{1 \le l \le L} \lambda_1[b_{j,c,l}[i]] \qquad (36)$$

where the selected estimate is the value $\lambda_1[b_{j,c,l_{\text{opt}}}[i]]$ which maximizes the likelihood and corresponds to the most likely bit. Based on the selected prior information $\lambda^p_1[b_{j,c,l_{\text{opt}}}[i]]$ and the trellis structure of the code, the MAP decoder processing the jth data stream and the lth branch computes the *a posteriori* LLR of each coded bit as described by

$$\Lambda_2[b_{j,c}[i]] = \log \frac{P[b_{j,c}[i] = +1|\lambda^p_1[b_{j,c,l_{\text{opt}}}[i]; \text{decoding}]}{P[b_{j,c}[i] = -1|\lambda^p_1[b_{j,c,l_{\text{opt}}}[i]; \text{decoding}]}$$

$$= \lambda_2[b_{j,c}[i]] + \lambda^p_1[b_{j,c,l_{\text{opt}}}[i]],$$

$$\text{for } j = 1, \dots, N_T, \ c = 1, \dots, C. \qquad (37)$$

The computational burden can be significantly reduced using the max-log approximation. From the above, it can be seen that the output of the MAP decoder is the sum of the prior information $\lambda^p_1[b_{j,c,l_{\text{opt}}}[i]]$ and the extrinsic information $\lambda_2[b_{j,c}[i]]$ produced by the MAP decoder. This extrinsic information is the information about the coded bit $b_{j,c}[i]$ obtained from the selected prior information about the other coded bits $\lambda^p_1[b_{j,c,l_{\text{opt}}}[k]]$, $j \ne i$ [44]. The MAP decoder also computes the *a posteriori* LLR of every information bit, which is used to make a decision on the decoded bit at the last iteration. After interleaving, the extrinsic information obtained by the MAP decoder $\lambda_2[b_{j,c}[i]]$ for $j = 1, \dots N_T, c = 1, \dots, C$ is fed back to the MB-MMSE-DF detector, as the prior information about the coded bits of all streams in the subsequent iteration. For the first iteration, $\lambda_1[b_{j,c}[i]]$ and $\lambda_2[b_{j,c,l_{\text{opt}}}[i]]$ are statistically independent and as the iterations are computed they become
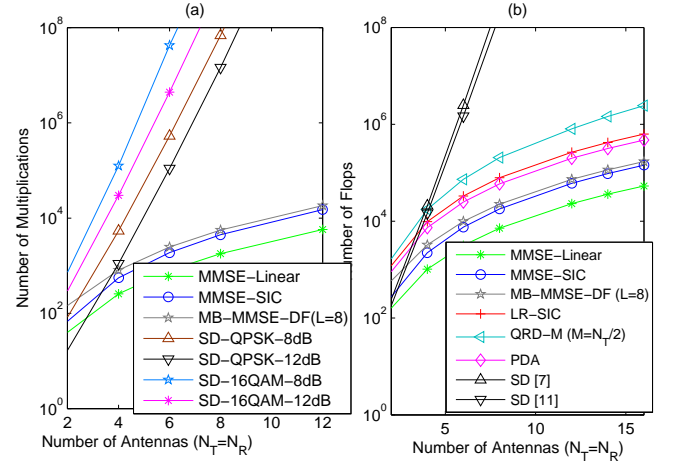

Fig. 3. Computational complexity of detection algorithms.

more correlated and the improvement due to each iteration is gradually reduced. A study of the proposed iterative MB-MMSE-DF detector has indicated that there is no performance gain when using more than 5 iterations.

VI. ANALYSIS OF THE MB-MMSE-DF ALGORITHM

In this section, the computational complexity required by the MB-MMSE-DF algorithm is evaluated and the diversity order achieved by the proposed MB-MMSE-DF detector is discussed.

A. Computational Complexity

The computational complexity of the MB-MMSE-DF detector can be exactly computed as a function of the number of receive antennas N_R, transmit antennas N_T and branches L, as depicted in Table III. This is in contrast to the SD and the LR-aided techniques, which require the use of bounds or the counting of floating point operations (flops). In this study of the computational cost of the MB-MMSE-DF and other techniques, two approaches to assess the complexity are employed, namely, the number of arithmetic operations such as multiplications and additions, and the number of flops computed by the Lightspeed toolbox [51]. [1] The complexity of the SD is associated with $M(\cdot)$, the Gamma function $\Gamma(\cdot)$, and the k-dimensional sphere radius $d_S D$, which is chosen as a scaled version of the variance of the noise [9]. The channel estimation with an RLS algorithm requires $N_R N_T^2 + 4N_T^2 + 2N_T N_R + 2N_T + 2$ multiplications and $N_R N_T^2 + 4N_T^2 - N_T$ additions.

An example of the computational cost of some detection algorithms is shown in Fig. 3, where the number of multiplications and flops per received vector $\boldsymbol{r}[i]$ are shown for the proposed MB-MMSE-DF and RLS algorithms, the bounds on the SD reported by [9] and the SD schemes of [7],[12], the complex LR-SIC of [23] with an MMSE filter, the PDA algorithm reported in [16] with $I = 5$ iterations, the linear

---

[1]According to the Lightspeed toolbox [51] the number of flops count as 2 for a complex addition and as 6 for a complex multiplication.

TABLE III
COMPUTATIONAL COMPLEXITY OF DETECTION ALGORITHMS PER RECEIVED VECTOR.

Algorithm	Number of operations per symbol	
	Additions	Multiplications
MB-MMSE-DF + RLS	$2N_R^2 + N_R N_T - 1$ $+L(3N_R N_T^2 + 2N_T^2)$ $-3N_R N_T + N_R - N_T$	$3N_R^2 + 2N_R N_T$ $+3N_R + 1$ $+L(5N_R N_T^2 + 2N_R)$
SIC + RLS [19]	$\frac{2}{3}N_R^3 + \frac{11}{2}N_R^2 + 4N_R$	$\frac{2}{3}N_R^3 + \frac{25}{2}N_R^2 + 3N_R$
Linear + RLS	$N_T(3N_R^2 + 2N_R - 1)$ $+2N_R N_T$	$N_T(3N_R^2 + 4N_R + 1)$
SD [9]	$\sum_{k=1}^{N_T} \frac{M(k+1)\pi^{k/2}}{\Gamma(k/2+1)} d_{\text{SD}}^k$ $+2N_T^2 - N_T + 2$	$\sum_{k=1}^{N_T} \frac{M(k)\pi^{k/2}}{\Gamma(k)/2+1)} d_{\text{SD}}^k$ $+2N_T^2$

detector and the SIC detector [19]. The complexity evaluated in terms of flops assumes 16-QAM modulation and $\text{SNR} = 8$ dB and includes the QRD-M detector [24], [25] with $M = 8$. The QRD-M algorithm is a breadth-first tree search algorithm, whereas the LSD is a depth-first tree search algorithm that can achieve the optimal performance. Differently from the QRD-M algorithm and the LSD algorithms, the proposed MB-MMSE-DF detector associates branches with different orderings and pairs of linear and feedback filters that only require one matrix inversion. The list of candidates in the MB-MMSE-DF algorithm is different because the candidates are generated by MMSE filtering and feedback cancellation with different orderings, while the list is generated from a tree search in the case of the QRD-M detector and the LSD algorithm. Moreover, the complexity of the MB-MMSE-DF detector only depends on the number of branches regardless of the constellation size and the signal-to-noise ratio (SNR), whereas the complexity of the QRD-M depends on the choice of M and the cost of LSD algorithms is dependent on the constellation size and the SNR. The curves of Fig. 3 indicate that the proposed MB-MMSE-DF and RLS algorithms have a complexity higher than the SIC [19] and significantly lower than the SD algorithms for $N_R = N_T > 4$ and the QRD-M algorithm. The MB-MMSE-DF detector also has a lower complexity than the LR-SIC and the PDA algorithms.

B. Diversity Order

The aim of this part is to examine the diversity order achieved by the MB-MMSE-DF detector. In the analysis, it is assumed that the data transmission is over a block fading channel, there is no error propagation due to interference cancellation and that the SNR is sufficiently high [52], [53] (in this case the MMSE and zero-forcing receive filters have a similar behavior). The diversity order [52], [53] is defined by

$$d \triangleq \lim_{\text{SNR} \to \infty} \frac{\log P_e(\text{SNR})}{\log(\text{SNR})}, \quad (38)$$

where P_e denotes the probability of error and $\text{SNR} = 10\log_{10} \frac{N_T \sigma_s^2}{RC\,\sigma^2}$ is the signal-to-noise ratio, R is the rate of the code and C is the number of bits required to map the constellation points. It is known that the diversity order is $d = N_R$ for ML receivers and $d = N_R - N_T + 1$ for receivers with SIC [52], [53]. Since for non-ergodic scenarios the error

probability is dominated by the outage probability [52], [53], the diversity order can be expressed as

$$d \triangleq \lim_{x \to \infty} \frac{\log Pr(R_{j,\text{span}\{1,2,\ldots,j-1,j+1,\ldots,N_T\}} \leq x)}{\log(x)}, \quad (39)$$

where $R_{j,\text{span}\{1,2,\ldots,j-1,j+1,\ldots,N_T\}} = R_{j,\text{span}\{\bar{j}\}}$ is the squared projection height from the jth column vector $\boldsymbol{h}_j$ of $\boldsymbol{H}$, i.e., $R_{j,\text{span}\{\bar{j}\}} = ||\boldsymbol{P}\boldsymbol{h}_j||^2$, where $\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{B}\boldsymbol{B}^H$ is the projection matrix onto the orthogonal space of $\text{span}\{\bar{j}\}$, and $\boldsymbol{B}$ is composed of any orthogonal basis of this subspace.

Theorem: The diversity order achieved by the MB-MMSE-DF detector is given by

$$d_{\text{MB}} \triangleq \lim_{x \to 0} \frac{\log Pr(R_{j,l_{j,\text{opt}},\text{span}\{\bar{l}_{j,\text{opt}}\}} \leq x)}{\log(x)} \quad (40)$$
$$= N_R - N_T + G,$$

where $1 \leq G \leq N_T$ is the number of interference free candidates among the L candidates for each stream.

Proof: In order to prove this theorem, it is necessary to make a few assumptions that are common to works that analyze the diversity order of detectors. The approach used to prove the theorem is based on induction and the inclusion of an increasing number of branches that correspond to extra degrees of freedom.

The first assumption is that for each data stream and branch there is an associated diversity order given by $d \triangleq \lim_{x\to\infty} \frac{\log Pr(R_{j,l,\text{span}\{\bar{j}\}} \leq x)}{\log(x)}$, as established in [52], [53] for a conventional receiver performing SIC in a MIMO system with N_T transmit and N_R receive antennas. Another assumption is that the ordering algorithm can exploit the multiple branches to move each data stream to the last position in the sequence of detection to obtain an interference free candidate for detection.

Starting from this point, the result can be extended by induction. By gradually adding branches with different orderings, the number of detection candidates available can be represented by the following sets

$$\begin{aligned} \mathcal{S}_1 &= \{R_{j,1,\text{span}\{\bar{j}\}}\}, \\ \mathcal{S}_2 &= \{R_{j,1,\text{span}\{\bar{j}\}}, R_{j,2,\text{span}\{\bar{j}\}}\}, \\ &\vdots \\ \mathcal{S}_L &= \{R_{j,1,\text{span}\{\bar{j}\}}, R_{j,2,\text{span}\{\bar{j}\}}, \ldots, R_{j,L,\text{span}\{\bar{j}\}}\}. \end{aligned} \quad (41)$$

Unlike a conventional receiver with SIC, the proposed MB-MMSE detector has at any given stage L alternatives to select

the candidate for detection. In fact, the detection of each stream involves the selection of the best out of L candidate symbols using the rule in (3). The number of degrees of freedom will depend on the quantities $R_{j,l_{j,\text{opt}},\text{span}\{\bar{l}_{j,\text{opt}}\}}$ and wether they correspond to interference free candidates.

By defining $1 \leq G \leq N_T$ as the number of interference free candidates and assuming that $L > G$ is sufficiently large to provide a sufficiently large number of interference free candidates for the jth stream, the diversity order associated with each of the above sets can be described by (42) where

$$l_{j,\text{opt}} = \arg \min_{1 \leq l_j \leq L} \text{IMMSE}(\sigma_s^2, \boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}) \qquad (43)$$

and $R_{j,l_{j,\text{opt}},\text{span}\{\bar{l}_{j,\text{opt}}\}}$ is the squared projection height resulting from the selection of the best out of the available candidates from the set $\mathcal{S}_l$ for the jth stream. If interference free candidates are gradually included in the sets and are selected by the above procedure, then the MB-MMSE-DF detector can obtain G interference free candidates resulting from $N_T - 1$ cancellations for any branch. Hence, the diversity order for each stream of the MB-MMSE-DF detector is given by

$$d_{\text{MB}} \triangleq \lim_{x \to 0} \frac{\log Pr(R_{j,l_{j,\text{opt}},\text{span}\{\bar{l}_{j,\text{opt}}\}} \leq x)}{\log(x)} \qquad (44)$$
$$= N_R - N_T + G.$$

This suggests that the key advantage of the MB-MMSE-DF detector is its ability to generate L candidates for each stream and select G interference free candidates. In practice, G will depend on the number of branches used, the ordering algorithm and the accuracy of the interference cancellation.

VII. SIMULATIONS

In this section, the bit error ratio (BER) performance of the MB-MMSE-DF and other relevant MIMO detection schemes is evaluated. The sphere decoder (SD) [12], the linear [13], the SIC [14], the QRD-M [24], [25] with $M = 8$, the PDA [26], [47] with $I = 5$ iterations, MMSE estimators and the proposed MB-MMSE-DF techniques without and with error propagation mitigation techniques are considered in the simulations. The lattice-reduction aided versions of the linear and the SIC detectors [23], which are denoted LR-MMSE-Linear and LR-MMSE-SIC, respectively, are also included in the study. The channel coefficients are either static and obtained from complex Gaussian random variables with zero mean and unit variance, or time-varying with the coefficients given by the Jakes model [54]. The modulation employed is either QPSK or 16-QAM. Both uncoded and coded systems are considered. For the coded systems and iterative detection and decoding, a non-recursive convolutional code with rate $R = 1/2$, constraint length 3, generator polynomial $g = [7\ 5]_{\text{oct}}$ and 5 decoding iterations is adopted. The numerical results are averaged over 10^6 runs, packets with $Q = 500$ symbols for uncoded systems and $Q = 1000$ coded symbols are employed and the signal-to-noise ratio (SNR) in dB is defined as SNR $= 10 \log_{10} \frac{N_T \sigma_s^2}{RC\ \sigma^2}$, where σ_s^2 is the variance of the symbols, σ^2 is the noise variance, $R < 1$ is the rate of the channel code and C is the number of bits used to represent the constellation.
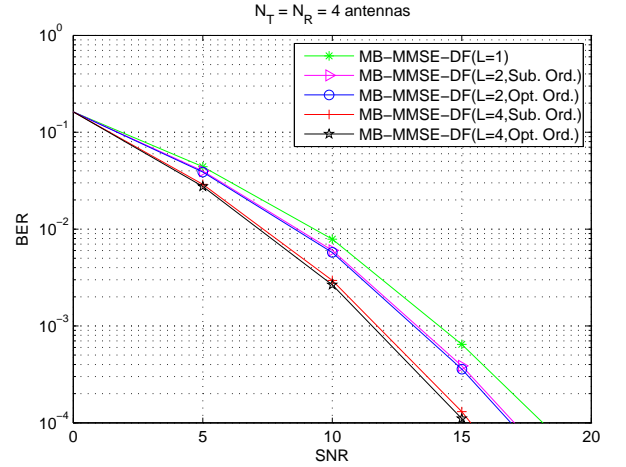

Fig. 4. BER performance of the optimal and the proposed suboptimal ordering algorithms .
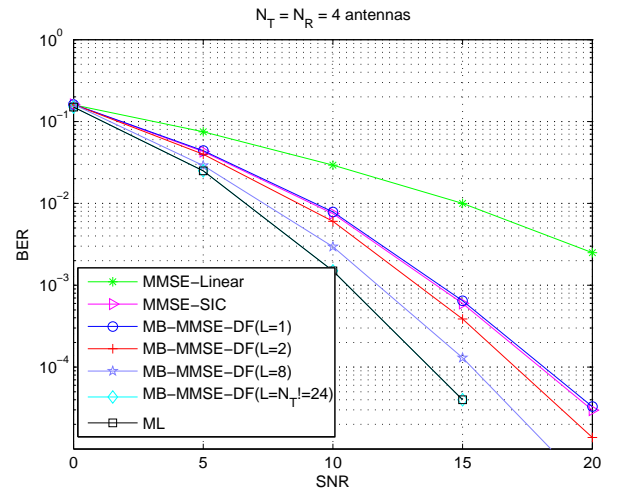

Fig. 5. BER performance of the detectors with perfect channel estimation for multiple branches L and QPSK.

In the first example, the ordering algorithms described in Section IV are assessed with the MB-MMSE-DF detector using $L = 1, 2,$ and 4 branches. A MIMO system with $N_T = N_R = 4$ antennas is considered with perfect channel estimation. The BER performance of the MB-MMSE-DF detector is evaluated with the optimal and the suboptimal ordering algorithms and the curves are shown in Fig. 4. The results show that the suboptimal ordering algorithm is able to approach the performance of the optimal ordering algorithm that performs an exhaustive search. In particular, the BER performance gap between the optimal and suboptimal ordering algorithms is small and this has also been observed for larger systems and a different number of branches L. For this reason, the suboptimal ordering algorithm has been adopted in the next examples.

The uncoded BER performance of the proposed MB-MMSE-DF detector is then considered in an example to evaluate the number of branches that should be used in the

$$d_{\mathrm{MB}}(\mathcal{S}_1) = \lim_{x \to 0} \frac{\log Pr(R_{j,l_{j,\mathrm{opt}},\mathrm{span}\{\bar{l}_{j,\mathrm{opt}}\}} \le x)}{\log(x)} = N_R - N_T + 1,$$

$$N_R - N_T + 1 \le d_{\mathrm{MB}}(\mathcal{S}_2) = \lim_{x \to 0} \frac{\log Pr(R_{j,l_{j,\mathrm{opt}},\mathrm{span}\{\bar{l}_{j,\mathrm{opt}}\}} \le x)}{\log(x)} \le N_R - N_T + 2,$$

$$\vdots$$

$$N_R - N_T + 1 \le d_{\mathrm{MB}}(\mathcal{S}_L) = \lim_{x \to 0} \frac{\log Pr(R_{j,l_{j,\mathrm{opt}},\mathrm{span}\{\bar{l}_{j,\mathrm{opt}}\}} \le x)}{\log(x)} \le N_R - N_T + G.$$
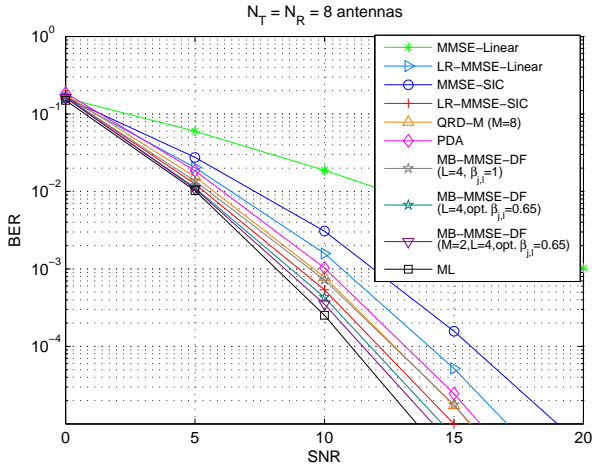
$$(42)$$

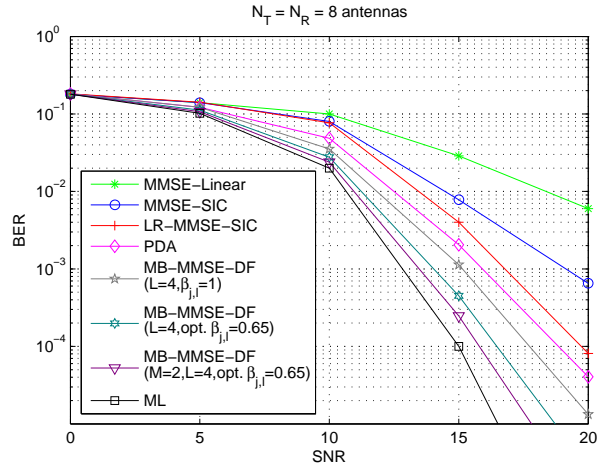Fig. 6. BER performance with adaptive estimation and QPSK modulation.

Fig. 7. BER performance with adaptive estimation and 16−QAM.

suboptimal ordering algorithm. It is also important to account for the impact of additional branches on the performance with perfect channel estimation for a MIMO system with $N_T = N_R = 4$ antennas. The proposed suboptimal ordering algorithm is compared against the optimal ordering approach described in Section IV that evaluates $N_T! = 24$ possible branches. The MB-MMSE-DF detector has been designed with $L = 2$ and 4 parallel branches and its BER performance against the SNR has been compared with those of the existing schemes, as depicted in Fig. 5. In fact, the MB-MMSE-DF detector is able to gradually approach the BER performance of the ML detector as the number of branches L is increased. Starting with L=1, the MB-MMSE-DF detector has a BER performance comparable with that of a standard MMSE-SIC detector. By increasing L the BER performance of MB-MMSE-DF gradually improves and gets within 1.5 dB of SNR for the same BER performance as the ML detector when L=8. Finally, MB-MMSE-DF obtains a performance that is comparable (the curves coincide) when L=24.

In the next experiments depicted in Figs. 6 and 7, the uncoded BER performance of the proposed MB-MMSE-DF detector is evaluated with $L = 4$, $N_T = N_R = 8$ antennas, QPSK and 16-QAM modulation, a block fading channel and adaptive estimation using the proposed RLS-type algorithm with $\hat{\boldsymbol{R}}^{-1}[0] = 10^{-2}\boldsymbol{I}$ and $\lambda = 0.998$. In the transmission, we assume packets with 500 symbols and employ a training sequence with $N_{Tr} = 50$ symbols to compute the channel and the receive filter coefficients. We include in the comparison

the linear and SIC detectors with RLS algorithms, the LR-MMSE-Linear and LR-MMSE-SIC detection schemes [23] using MMSE filters, the QRD-M technique of [24], [25], the PDA algorithm of [26] and the SD of [12] to compute the ML solution, which employ the RLS algorithm to estimate the channels.

The results depicted in Figs. 6 and 7 for QPSK and 16-QAM, respectively, show that the proposed MB-MMSE-DF detector achieves a performance which is close to the ML solutions implemented with the SD and outperforms the linear, the SIC, the LR-MMSE-Linear, LR-MMSE-SIC, the QRD-M and the PDA detectors by a significant margin. In particular, the proposed MB-MMSE-DF detector with $L = 4$ without error propagation mitigation ($\beta_{j,l} = \gamma_{j,l} = 1$) has a comparable performance to the PDA and the LR-MMSE-SIC detectors, whereas the MB-MMSE-DF scheme with an optimized value of $\beta_{j,l} = 0.65$ outperforms the PDA and the LR-MMSE-SIC schemes. The MB-MMSE-DF scheme with $M = 2$ stages and $L = 4$ significantly outperforms the QRD-M, the PDA and the LR-MMSE-SIC algorithms and achieves a performance within 1 dB from the ML solution, while requiring a cost comparable to the SIC with the RLS algorithm.

In the next two examples, the uncoded and coded BER performances of the detectors are assessed for systems with 16-QAM modulation and time-varying channels. The channel coefficients in these examples change every received vector according to the Jakes model [54] and the results are
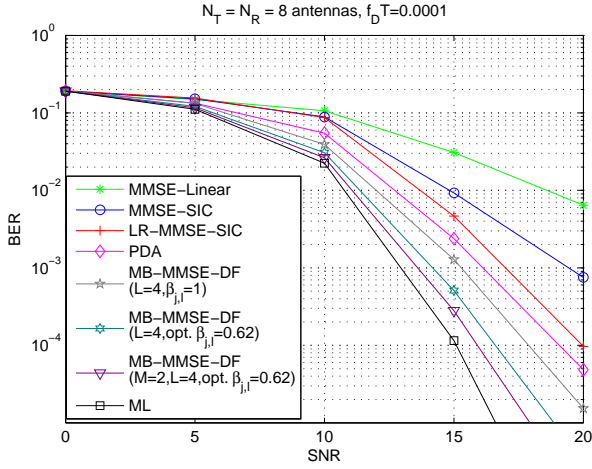
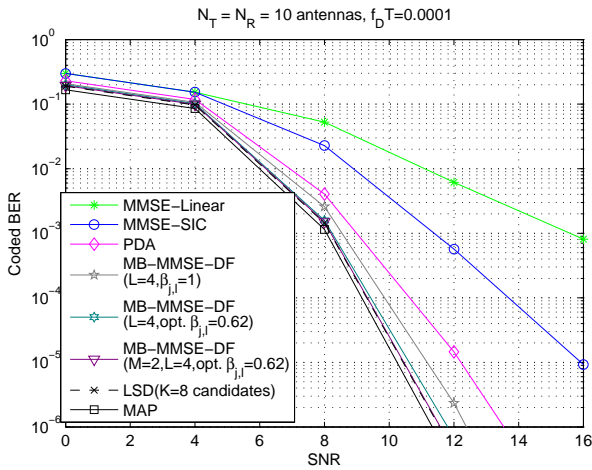Fig. 8. BER performance with adaptive estimation and $16-$QAM in time-varying channels with $f_D T = 10^{-4}$.

Fig. 9. Coded BER performance with adaptive estimation and $16-$QAM in time-varying channels with $f_D T = 10^{-4}$ and 5 decoding iterations.

shown in terms of the normalized Doppler frequency $f_D T$ (cycles/symbol), where f_D is the maximum Doppler shift and T is the symbol interval. For the data transmission, packets with $Q = 500$ symbols are used for the uncoded system, with $Q = 1000$ coded symbols for the convolutionally coded system with 5 decoding iterations, and a training sequence with $N_{Tr} = 50$ symbols is employed to compute the channel and receive filter coefficients. After the training sequence, the receivers are switched to decision-directed mode and the parameters are tracked with RLS-type algorithms.

The uncoded BER results illustrated in Fig. 8 show similar results to that of Fig. 7 but with a slight performance degradation due to the time-varying nature of the channel. The coded BER performance illustrated in Fig. 9 indicates that the proposed iterative MB-MMSE-DF detector with an optimized value of $\beta_{j,l} = 0.62$, $L = 4$ and $M = 2$ has a performance that is very close to the optimal MAP detector and is comparable to the list SD (LSD) with $K = 8$ candidates,
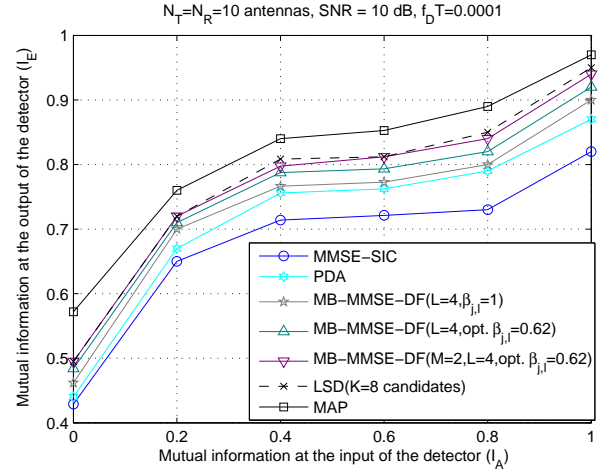
which corresponds to the SD of [12] with LLR processing. The proposed iterative MB-MMSE-DF detector has a gain of up to 2 dB over the PDA detector and of up to 5 dB over the conventional SIC with iterative processing for the same coded BER, while the computational cost of MB-MMSE-DF is significantly lower than the PDA technique and slightly higher than the SIC algorithm.

In Fig. 10, the soft input and output behavior of the detection algorithms is described through the use of the extrinsic information transfer (EXIT) chart [55] analysis. In this plot, 16-QAM modulation with a 10×10 MIMO system are considered. The quantities I_A and I_E represent the mutual information at the input and at the output of the detectors analyzed. The proposed MB-MMSE-DF detector is able to achieve a higher capacity compared to the other suboptimal algorithms considered and to follow closely the trajectory of the LSD and MAP algorithms. Specifically, with the increased number of branches, more tentative decisions or candidates are included in the search space for the solution and this allows the proposed MB-MMSE-DF detector to approach the performance of the MAP algorithm.

VIII. CONCLUDING REMARKS

This work has proposed and investigated MB-MMSE-DF detection algorithms for large MIMO systems using spatial multiplexing. Constrained MMSE filters designed with constraints on the shape and magnitude of the feedback filters have been presented for the MB-MMSE-DF detector and it has been shown that the proposed design does not require a significant additional complexity over the conventional MMSE-DF detector. Optimal and sub-optimal ordering algorithms have also been presented for the MB-MMSE-DF detector. An adaptive version of the MB-MMSE-DF detector has been developed with an RLS-type algorithm for estimating the parameters of the filters when the channel is time-varying. A soft-output version of the MB-MMSE DF detector has also been proposed as a component of an iterative detection and decoding receiver structure. The results have shown that the proposed MB-MMSE-DF detector achieves a performance

Fig. 10. EXIT chart for the analyzed detectors with $16-$QAM.

superior to some existing suboptimal detectors and close to the ML detector, while requiring significantly lower complexity.

APPENDIX

In order to derive the MMSE receive filters resulting from the constrained optimization problem presented in (6), the method of Lagrange multipliers is used which results in the following unconstrained cost function (45) where $\boldsymbol{\alpha}_{j,l}$ is an $N_T \times 1$ vector of Lagrange multipliers and $\mu_{j,l}$ is a scalar Lagrange multiplier.

By computing the gradient terms of (45) with respect to $\boldsymbol{w}_{j,l}^*$ and equating them to zero, we have

$$\nabla \mathcal{L}(\boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{\alpha}_{j,l}, \mu_{j,l})_{\boldsymbol{w}_{j,l}^*} = \\ E[-\boldsymbol{r}[i](s_j[i] - \boldsymbol{w}_{j,l}^H \boldsymbol{r}[i] + \boldsymbol{f}_{j,l}^H \hat{\boldsymbol{s}}_l[i])^*] = \mathbf{0} \quad (46)$$

By further manipulating the terms in the above equation, we arrive at the expression obtained in (7)

$$\boldsymbol{w}_{j,l}^{\mathrm{MMSE}} = \boldsymbol{R}^{-1}(\boldsymbol{p}_j + \boldsymbol{Q}\boldsymbol{f}_{j,l}), \quad (47)$$

where $\boldsymbol{R} = E[\boldsymbol{r}[i]\boldsymbol{r}^H[i]]$ is the $N_R \times N_R$ covariance matrix of the received data, $\boldsymbol{p}_j = E[\boldsymbol{r}[i]s_j^*[i]]$ is the $N_R \times 1$ cross-correlation vector and $\boldsymbol{Q} = E[\boldsymbol{r}[i]\hat{\boldsymbol{s}}_l^H[i]]$ is a $N_R \times N_T$ cross-correlation matrix.

By calculating the gradient terms of (45) with respect to $\boldsymbol{f}_{j,l}^*$ and equating them to zero, we have

$$\nabla \mathcal{L}(\boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{\alpha}_{j,l}, \mu_{j,l})_{\boldsymbol{f}_{j,l}^*} = \\ E[\hat{\boldsymbol{s}}_l[i](s_j[i] - \boldsymbol{w}_{j,l}^H \boldsymbol{r}[i] + \boldsymbol{f}_{j,l}^H \hat{\boldsymbol{s}}_l[i])^*] + 2\boldsymbol{S}_{j,l}\boldsymbol{\alpha}_{j,l} + 2\boldsymbol{f}_{j,l}\mu_{j,l} = \mathbf{0} \quad (48)$$

Using the above equation and with further manipulations, we obtain

$$\boldsymbol{f}_{j,l}^{\mathrm{MMSE}} = \frac{\beta_{j,l}}{\sigma_s^2}(\boldsymbol{Q}^H \boldsymbol{w}_{j,l} - \boldsymbol{t}_j - 2\boldsymbol{S}_{j,l}^H \boldsymbol{\alpha}_{j,l}). \quad (49)$$

where the term $\beta_{j,l} = (1 - 2\mu_{j,l})^{-1}$ with the Lagrange multiplier $\mu_{j,l}$ is responsible for the mitigation of the error propagation and is a parameter to be adjusted, and $E[\hat{\boldsymbol{s}}[i]\boldsymbol{s}^H[i]] = \sigma_s^2 \boldsymbol{I}$ since it is assumed that $\hat{\boldsymbol{s}}[i]$ has independent and identically distributed entries. The above expression describes the relationship between the feedback filters $\boldsymbol{f}_{j,l}^{\mathrm{opt}}$, the feedforward filter $\boldsymbol{w}_{j,l}^{\mathrm{opt}}$ and the quantities $\boldsymbol{Q}$, $\boldsymbol{t}_j = E[\hat{\boldsymbol{s}}_l[i]s_j[i]]$, and the Lagrange multipliers $\mu_{j,l}$ and $\boldsymbol{\alpha}_{j,l}$.

The expression for the Lagrange multiplier $\boldsymbol{\alpha}_{j,l}$ can be obtained by computing the gradient terms of (45) with respect to $\boldsymbol{\alpha}_{j,l}$ and equating them to zero, which results in

$$\nabla \mathcal{L}(\boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{\alpha}_{j,l}, \mu_{j,l})_{\boldsymbol{\alpha}_{j,l}} = \boldsymbol{S}_{j,l}\boldsymbol{f}_{j,l} = \mathbf{0} \quad (50)$$

By substituting (49) into the above expression and solving for $\boldsymbol{\alpha}_{j,l}$, we have

$$\boldsymbol{\alpha}_{j,l} = (\boldsymbol{S}_{j,l}\boldsymbol{S}_{j,l}^H)^{-1}(\boldsymbol{S}_{j,l}^H \boldsymbol{Q}^H \boldsymbol{w}_{j,l} - \boldsymbol{S}_{j,l}\boldsymbol{t}_j)/2, \quad (51)$$

It turns out that there is no closed form solution for the term $\beta_{j,l}$, which is a function of the Lagrange multiplier $\mu_{j,l}$. This happens because its evaluation leads to a quadratic function of the feedback filter $\boldsymbol{f}_{j,l}$ that is quite involved. For this reason, we employ an approach that computes $\beta_{j,l}$ numerically. By inserting the expression for $\boldsymbol{\alpha}_{j,l}$ into (49), we arrive at

$$\boldsymbol{f}_{j,l}^{\mathrm{MMSE}} = \\ \frac{\beta_{j,l}}{\sigma_s^2}\Big(\boldsymbol{Q}^H \boldsymbol{w}_{j,l} - \boldsymbol{t}_j - \boldsymbol{S}_{j,l}^H(\boldsymbol{S}_{j,l}\boldsymbol{S}_{j,l}^H)^{-1}(\boldsymbol{S}_{j,l}^H \boldsymbol{Q}^H \boldsymbol{w}_{j,l} - \boldsymbol{S}_{j,l}\boldsymbol{t}_j)\Big), \quad (52)$$

where $\boldsymbol{f}_{j,l}^{\mathrm{MMSE}}$ is a function of the Lagrange multiplier $\mu_{j,l}$. By defining $\boldsymbol{\Pi}_{j,l} = \boldsymbol{I} - \boldsymbol{S}_{j,l}^H(\boldsymbol{S}_{j,l}^H \boldsymbol{S}_{j,l})^{-1}\boldsymbol{S}_{j,l}$ as a projection matrix that ensures the shape constraint $\boldsymbol{S}_{j,l}$, the above expression can be written as (8) in the compact form

$$\boldsymbol{f}_{j,l}^{\mathrm{MMSE}} = \frac{\beta_{j,l}}{\sigma_s^2}\boldsymbol{\Pi}_{j,l}(\boldsymbol{Q}^H \boldsymbol{w}_{j,l} - \boldsymbol{t}_j), \quad (53)$$

Now if we substitute the above expression into (47) and further manipulate the expressions, we obtain

$$\boldsymbol{w}_{j,l}^{\mathrm{MMSE}} = (\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H)^{-1}(\boldsymbol{p}_j - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{t}_j). \quad (54)$$

Substituting the above expression into (53), we obtain (55) where the above expressions for the receive filters $\boldsymbol{w}_{j,l}$ and $\boldsymbol{f}_{j,l}$ only depend on the statistical quantities $\boldsymbol{R}$, $\boldsymbol{Q}$, $\boldsymbol{t}_j$ and $\boldsymbol{p}_j$ and the parameters $\beta_{j,l}$ and $\boldsymbol{\Pi}_{j,l}$. Nevertheless, these expressions have an inconvenient form for practical use as they require multiple matrix inversions for the computation of the receiver filters for each data stream j and branch l. To circumvent this drawback, the use of an alternating strategy with (47) and (52) is employed as it allows a designer to compute only one matrix inversion ($\boldsymbol{R}^{-1}$) and all the receive filters with a reduced number of extra multiplications and additions.

The expressions obtained so far can be simplified by evaluating some of the key statistical quantities such as $\boldsymbol{R}$, $\boldsymbol{Q}$, $\boldsymbol{t}_j$ and $\boldsymbol{p}_j$ and replacing them in the formulas for the receive filters. Using the fact that the quantity $\boldsymbol{t}_j = \mathbf{0}$ for interference cancellation, $\boldsymbol{v}_{j,l} = \mathbf{0}$, and assuming perfect feedback ($\boldsymbol{s} = \hat{\boldsymbol{s}}$) we have

$$\boldsymbol{R} = \sigma_s^2 \boldsymbol{H}\boldsymbol{H}^H + \sigma_n^2 \boldsymbol{I}, \boldsymbol{Q} = \sigma_s^2 \boldsymbol{H}, \boldsymbol{p}_j = \sigma_s^2 \boldsymbol{H}\boldsymbol{\delta}_j, \quad (56)$$

where $\boldsymbol{\delta}_j = [\underbrace{0\ldots0}_{j-1} \; 1 \; \underbrace{0\ldots0}_{N_T-j-2}]^T$ is a $N_T \times 1$ vector with a one in the jth element and zeros elsewhere. Substituting $\boldsymbol{t}_j = \mathbf{0}$ and (56) into (47) and (52) we arrive at (12) and (13), respectively.

The MMSE associated with the filters $\boldsymbol{w}_{j,l}$ and $\boldsymbol{f}_{j,l}$ and the statistics of the data symbols $s_j[i]$ is given by (57). By substituting $\boldsymbol{t}_j = \mathbf{0}$, the quantities in (56) and the expressions in (54) and (55), the MMSE becomes (58). For a given ordering and branch l, the sum of the MMSE in (58) over the N_T data streams is equivalent to the MMSE achieved by a conventional MMSE-DF receiver (C-MMSE-DF) and is given by

$$\mathrm{SMMSE} = \sum_{j=1}^{N_T} MMSE_j \quad (59)$$

An instantaneous MMSE metric for the selection of the best branch for each received vector can be obtained by removing the expected value from the expression in (57) and considering each received data vector $\boldsymbol{r}[i]$, which results in (60). The expression above suggests that in order to obtain an instantaneous MMSE metric, the MB-MMSE-DF detector needs to compute all the terms. However, it is possible to obtain an effective and yet more efficient expression by inspecting the terms in the first line of (58) and retaining the corresponding instantaneous values, which results in

$$\mathrm{IMMSE}(s_j[i], \boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{r}[i]) \approx \\ |s_j[i]|^2 - \boldsymbol{w}_{j,l}^{H,\;\mathrm{MMSE}}\hat{\boldsymbol{R}}\boldsymbol{w}_{j,l}^{\mathrm{MMSE}} + \boldsymbol{f}_{j,l}^{H,\;\mathrm{MMSE}}\boldsymbol{s}_l\boldsymbol{s}_l^H \boldsymbol{f}_{j,l}^{\mathrm{MMSE}}. \quad (61)$$

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{\alpha}_{j,l}, \mu_{j,l}) = {} & E[|s_j[i] - \boldsymbol{w}_{j,l}^H \boldsymbol{r}[i] + \boldsymbol{f}_{j,l}^H \hat{\boldsymbol{s}}_l[i]|^2] + 2\Re[(\boldsymbol{S}_{j,l}\boldsymbol{f}_{j,l})^H \boldsymbol{\alpha}_{j,l}] \\
& + 2\Re[(\boldsymbol{f}_{j,l}^H \boldsymbol{f}_{j,l} - \gamma_{j,l}\boldsymbol{f}_{j,l}^{H,\ c}\boldsymbol{f}_{j,l}^c)^H \mu_{j,l}], \quad j = 1, \ldots, N_T,\ l = 1, \ldots, L,
\end{aligned}
\tag{45}
$$

$$
\boldsymbol{f}_{j,l}^{\mathrm{MMSE}} = \frac{\beta_{j,l}}{\sigma_s^2} \boldsymbol{\Pi}_{j,l}\Big[\boldsymbol{Q}^H(\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H)^{-1}\big(\boldsymbol{p}_j - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{t}_j\big) - \boldsymbol{t}_j\Big],
\tag{55}
$$

$$
\begin{aligned}
\underbrace{\mathrm{MMSE}(s_j[i], \boldsymbol{w}_{j,l}^{\mathrm{MMSE}}, \boldsymbol{f}_{j,l}^{\mathrm{MMSE}})}_{\mathrm{MMSE}_j} = {} & E[|s_j[i] - \boldsymbol{w}_{j,l}^H \boldsymbol{r}[i] + \boldsymbol{f}_{j,l}^H \boldsymbol{s}_l[i]|^2] \\
= {} & \underbrace{E[|s_j[i]|^2]}_{\sigma_s^2} - \boldsymbol{w}_{j,l}^H \underbrace{E[\boldsymbol{r}[i]s_j^*[i]]}_{\boldsymbol{p}_j} - \underbrace{E[\boldsymbol{r}^H[i]s_j[i]]}_{\boldsymbol{p}_j^H}\boldsymbol{w}_{j,l} + \boldsymbol{w}_{j,l}^H \underbrace{E[\boldsymbol{r}[i]\boldsymbol{r}^H[i]]}_{\boldsymbol{R}}\boldsymbol{w}_{j,l} \\
& - \boldsymbol{w}_{j,l}^H \underbrace{E[\boldsymbol{r}[i]\boldsymbol{s}_l^H[i]]}_{\boldsymbol{Q}}\boldsymbol{f}_{j,l} - \boldsymbol{f}_{j,l}^H \boldsymbol{Q}^H \boldsymbol{w}_{j,l} + \boldsymbol{f}_{j,l}^H \underbrace{E[\boldsymbol{s}_l s_j^*[i]]}_{\boldsymbol{t}_j} + \boldsymbol{t}_j^H \boldsymbol{f}_{j,l} + \boldsymbol{f}_{j,l}^H \underbrace{E[\boldsymbol{s}_l \boldsymbol{s}_l^H]}_{\boldsymbol{I}}\boldsymbol{f}_{j,l} \\
= {} & \sigma_s^2 - \boldsymbol{w}_{j,l}^H \boldsymbol{p}_j - \boldsymbol{p}_j^H \boldsymbol{w}_{j,l} + \boldsymbol{w}_{j,l}^H \boldsymbol{R}\boldsymbol{w}_{j,l} - \boldsymbol{w}_{j,l}^H \boldsymbol{Q}\boldsymbol{f}_{j,l} - \boldsymbol{f}_{j,l}^H \boldsymbol{Q}^H \boldsymbol{w}_{j,l} + \boldsymbol{f}_{j,l}^H \boldsymbol{t}_j + \boldsymbol{t}_j^H \boldsymbol{f}_{j,l} + \boldsymbol{f}_{j,l}^H \boldsymbol{f}_{j,l} \\
= {} & \sigma_s^2 - \boldsymbol{w}_{j,l}^H \boldsymbol{p}_j - \boldsymbol{p}_j^H \boldsymbol{w}_{j,l} + \underbrace{(\boldsymbol{f}_{j,l}^H \boldsymbol{Q}^H + \boldsymbol{p}_j^H)}_{\boldsymbol{w}_{j,l}}\boldsymbol{R}^{-1}\boldsymbol{w}_{j,l} \\
& - \boldsymbol{w}_{j,l}^H \boldsymbol{Q}\boldsymbol{f}_{j,l} - \boldsymbol{f}_{j,l}^H \boldsymbol{Q}^H \boldsymbol{w}_{j,l} + \boldsymbol{f}_{j,l}^H \boldsymbol{t}_j + \boldsymbol{t}_j^H \boldsymbol{f}_{j,l} + \boldsymbol{f}_{j,l}^H \boldsymbol{f}_{j,l} \\
= {} & \sigma_s^2 - \boldsymbol{w}_{j,l}^H \boldsymbol{R}\boldsymbol{w}_{j,l} + \boldsymbol{t}_j^H \boldsymbol{f}_{j,l} + \boldsymbol{f}_{j,l}^H \boldsymbol{t}_j + \boldsymbol{f}_{j,l}^H \boldsymbol{f}_{j,l}.
\end{aligned}
\tag{57}
$$

$$
\begin{aligned}
\mathrm{MMSE}_j = {} & \sigma_s^2 - \boldsymbol{w}_{j,l}^{H,\ \mathrm{MMSE}}\boldsymbol{R}\boldsymbol{w}_{j,l}^{\mathrm{MMSE}} + \boldsymbol{f}_{j,l}^{H,\ \mathrm{MMSE}}\boldsymbol{f}_{j,l}^{\mathrm{opt}} \\
= {} & \sigma_s^2 - \boldsymbol{p}_j^H(\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H)^{-1}\boldsymbol{R}(\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H)^{-1}\boldsymbol{p}_j \\
& + \beta_{j,l}^2 \boldsymbol{p}_j^H(\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H)^{-1}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}^H\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H(\boldsymbol{R} - \beta_{j,l}\boldsymbol{Q}\boldsymbol{\Pi}_{j,l}\boldsymbol{Q}^H)^{-1}\boldsymbol{p}_j.
\end{aligned}
\tag{58}
$$

The expression in (61) has been tested and compared with (60), and the results indicate an equivalent performance of the two expressions. Due to the smaller number of terms, the expression in (61) has been adopted for the operation of the MB-MMSE-DF detector.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas", *Wireless Pers. Commun.*, vol. 6, pp. 311335, Mar. 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels", *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, 1999.

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

[4] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1456-1467, July 1999.

[5] S. Verdu, *Multiuser Detection*, Cambridge, 1998.

[6] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels", *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, July 1999.

[7] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2389–2402, Oct. 2003.

[8] H. Vikalo, B. Hassibi, and T. Kailath, Iterative decoding for MIMO channels via modified sphere decoding," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 2299-2311, Nov. 2004.

[9] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm: Part I, the expected complexity", *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2806-2818, Aug. 2005.

[10] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-Best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.

[11] C. Studer, A. Burg, and H. Bolcskei, "Soft-output sphere decoding: algorithms and VLSI implementation," IEEE J. Sel. Areas Commun., vol. 26, pp. 290-300, Feb. 2008.

[12] B. Shim and I. Kang, "On further reduction of complexity in tree pruning based sphere search," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 417–422, Feb. 2010.

[13] A. Duel-Hallen, "Equalizers for Multiple Input Multiple Output Channels and PAM Systems with Cyclostationary Input Sequences," *IEEE J. Sel. Areas Commun.*, vol. 10, pp. 630-639, Apr., 1992.

[14] G. D. Golden, C. J. Foschini, R. A. Valenzuela and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture", *Electron. Lett.*, vol. 35, no. 1, Jan. 1999.

[15] J. Benesty, Y. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Trans. Signal Process.*, vol. 51, pp. 1722–1730, July 2003.

[16] Y. Shang and X.-G. Xia, "An improved fast recursive algorithm for V-BLAST with optimal ordered detections," in *Proc. IEEE ICC 2008*, Beijing, China, May 2008, pp. 756–760.

[17] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2921-2936, Oct., 2000.

[18] J. H. Choi, H. Y. Yu, Y. H. Lee, "Adaptive MIMO decision feedback equalization for receivers with time-varying channels", *IEEE Trans. Signal Proc.*, 2005, 53, no. 11, pp. 4295-4303.

[19] A. Rontogiannis, V. Kekatos, and K. Berberidis, "A square-root adaptive V-BLAST algorithm for fast time-varying MIMO channels," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 265-268, May 2006.

[20] R. Fa, R. C. de Lamare, "Multi-branch successive interference cancellation for MIMO spatial multiplexing systems", *IET Commun.*, vol. 5, no. 4, pp. 484-494, Mar. 2011.

$$\text{IMMSE}(s_j[i], \boldsymbol{w}_{j,l}, \boldsymbol{f}_{j,l}, \boldsymbol{r}[i]) = |s_j[i] - \boldsymbol{w}_{j,l}^H \boldsymbol{r}[i] + \boldsymbol{f}_{j,l}^H \boldsymbol{s}_l[i]|^2 = |s_j[i]|^2 - \boldsymbol{w}_{j,l}^H \underbrace{\boldsymbol{r}[i]s_j^*[i]}_{\hat{\boldsymbol{p}}_j} - \underbrace{\boldsymbol{r}^H[i]s_j[i]}_{\hat{\boldsymbol{p}}_j^H} \boldsymbol{w}_{j,l} + \boldsymbol{w}_{j,l}^H \underbrace{\boldsymbol{r}[i]\boldsymbol{r}^H[i]}_{\hat{\boldsymbol{R}}} \boldsymbol{w}_{j,l}$$

$$- \boldsymbol{w}_{j,l}^H \underbrace{\boldsymbol{r}[i]\boldsymbol{s}_l^H[i]}_{\hat{\boldsymbol{Q}}} \boldsymbol{f}_{j,l} - \boldsymbol{f}_{j,l}^H \hat{\boldsymbol{Q}}^H \boldsymbol{w}_{j,l} + \boldsymbol{f}_{j,l}^H \underbrace{\boldsymbol{s}_l s_j^*[i]}_{\hat{\boldsymbol{t}}_j} + \boldsymbol{t}_j^H \boldsymbol{f}_{j,l} + \boldsymbol{f}_{j,l}^H \boldsymbol{s}_l \boldsymbol{s}_l^H \boldsymbol{f}_{j,l}.$$

$$(60)$$

[21] P. Li, R. C. de Lamare and R. Fa, "Multiple feedback successive interference cancellation detection for multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2434-2439, Aug. 2011.

[22] C. Windpassinger, L. Lampe, R.F.H. Fischer, T.A Hehn, "A performance study of MIMO detectors," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, Aug. 2006, pp. 2004-2008.

[23] Y. H. Gan, C. Ling, and W. H. Mow, "Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection," *IEEE Trans. Signal Process.*, vol. 56, no. 7, July 2009.

[24] K. J. Kim, J. Yue, R. A. Iltis, and J. D. Gibson, "A QRD-M/Kalman filter-based detection and channel estimation algorithm for MIMO-OFDM systems", *IEEE Trans. Wireless Commun.*, vol. 4, pp. 710-721, Mar. 2005.

[25] K. J. Kim, T. Reid and R. A. Iltis, "Iterative soft-QRD-M for turbo coded MIMO-OFDM systems", *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1043-1046, July 2008.

[26] Y. Jia, C. M. Vithanage, C. Andrieu, and R. J. Piechocki, "Probabilistic data association for symbol detection in MIMO systems," *Electron. Lett.*, vol. 42, no. 1, pp. 38–40, Jan. 2006.

[27] S. Yang, T. Lv, R. Maunder, and L. Hanzo, "Unified bit-based probabilistic data association aided MIMO detection for high-order QAM constellations", *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 981-991, 2011.

[28] M. K. Varanasi, "Decision feedback multiuser detection: A systematic approach," *IEEE Trans. Inf. Theory*, vol. 45, pp. 219-240, Jan. 1999.

[29] J. F. Rößler and J. B. Huber, "Iterative soft decision interference cancellation receivers for DS-CDMA downlink employing 4QAM and 16QAM," in *Proc. 36th Asilomar Conf. Signal, Syst., Comput.*, Pacific Grove, CA, Nov. 2002.

[30] J. Luo, K. R. Pattipati, P. K. Willet and F. Hasegawa, "Optimal user ordering and time labeling for ideal decision feedback detection in asynchronous CDMA", *IEEE Trans. Commun.*, vol. 51, no. 11, Nov., 2003.

[31] G. Woodward, R. Ratasuk, M. L. Honig and P. Rapajic, "Minimum mean-squared error multiuser decision-feedback detectors for DS-CDMA," *IEEE Trans. Commun.*, vol. 50, no. 12, Dec., 2002.

[32] R.C. de Lamare, R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels", *IEEE Commun. Lett.*, vol. 7, no. 2, Feb. 2003, pp. 73-75.

[33] F. Cao, J. Li, and J. Yang, "On the relation between PDA and MMSE-ISDIC," *IEEE Signal Process. Lett.*, vol. 14, no. 9, Sept. 2007.

[34] R.C. de Lamare, R. Sampaio-Neto, "Minimum mean-squared error iterative successive parallel arbitrated decision feedback detectors for DS-CDMA systems", *IEEE Trans. Commun.*, vol. 56, no. 5, May 2008, pp. 778-789.

[35] Y. Cai and R. C. de Lamare, "Adaptive space-time decision feedback detectors with multiple feedback cancellation", *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, Oct. 2009, pp. 4129-4140.

[36] R. C. de Lamare and D. Le Ruyet, "Multi-branch MMSE decision feedback detection algorithms with error propagation mitigation for MIMO systems", *Proc. IEEE International Conf. Acoustics Speech Signal Process.*, Mar. 2010.

[37] P. Li and R. C. de Lamare, "Adaptive decision-feedback detection with constellation constraints for MIMO systems", *IEEE Trans. Veh. Technol.*, vol. 61, no. 2, pp. 853-859, 2012.

[38] M. Reuter, J.C. Allen, J. R. Zeidler, R. C. North, "Mitigating error propagation effects in a decision feedback equalizer", *IEEE Trans. Commun.*, vol. 49, no. 11, Nov. 2001, pp. 2028-2041.

[39] R.C. de Lamare, R. Sampaio-Neto, A. Hjorungnes, "Joint iterative interference cancellation and parameter estimation for cdma systems", *IEEE Commun. Lett.*, vol. 11, no. 12, Dec. 2007, pp. 916-918.

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

[41] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank adaptive filtering based on joint iterative optimization of adaptive filters, " *IEEE Signal Proc. Lett.*, vol. 14, no. 12, Dec. 2007, pp. 980-983.

[42] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering," *IEEE Trans. Signal Process.*, vol. 57, no. 7, July 2009, pp. 2503-2514.

[43] R.C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2482-2494, July 2011.

[44] 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.

[45] M. Tuchler, A. Singer, and R. Koetter, "Minimum mean square error equalization using a priori information," *IEEE Trans. Signal Process.*, vol. 50, pp. 673-683, Mar. 2002.

[46] B. Hochwald and S. ten Brink, "Achieving near-capacity on a mutliple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389-399, Mar. 2003.

[47] H. Lee, B. Lee, and I. Lee, "Iterative detection and decoding with an improved V-BLAST for MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 504-513, Mar. 2006.

[48] X. Yuan, Q. Guo, X. Wang, and Li Ping, "Evolution analysis of low-cost iterative equalization in coded linear systems with cyclic prefixes," *IEEE J. Sel. Areas Commun. (JSAC)*, vol. 26, no. 2, pp. 301-310, Feb. 2008.

[49] J. W. Choi, A. C. Singer, J Lee, N. I. Cho, "Improved linear soft-input soft-output detection via soft feedback successive interference cancellation," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 986-996, Mar. 2010.

[50] T. Hwang, Y. Kim, and H. Park, "Energy spreading transform approach to achieve full diversity and full rate for MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 12, Dec. 2012.

[51] T. Minka, "The Lightspeed Matlab toolbox, efficient operations for Matlab programming, version 2.2", 17 Dec. 2007. Microsoft Corp.

[52] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels", *IEEE Trans. Inf. Theory*, vol. 49, no. 5, May 2003.

[53] H. Zhang, H. Dai and B. L. Hughes, "Analysis of the diversity-multiplexing tradeoff for ordered MIMO SIC receivers," *IEEE Trans. Commun.*, vol. 57, no. 1, Jan. 2009.

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

[55] S. ten Brink, "Convergence behaviour of iteratively decoded parallel concatenated codes", *IEEE Trans. Commun.*, vol. 49, Oct. 2001.

Rodrigo C. de Lamare (S'99 - M'04 - SM'10) received the Diploma in electronic engineering from the Federal University of Rio de Janeiro (UFRJ) in 1998 and the M.Sc. and PhD degrees, both in electrical engineering, from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio) in 2001 and 2004, respectively. Since January 2006, he has been with the Communications Research Group, Department of Electronics, University of York, where he is currently a Reader. Since April 2013, he has also been a Professor with PUC-RIO. His research interests lie in communications and signal processing, areas in which he has published around 300 papers in refereed journals and conferences. Dr. de Lamare serves as associate editor for the EURASIP Journal on Wireless Communications and Networking and the IEEE Signal Processing Letters. He is a Senior Member of the IEEE and has served as the General Chair of the 7th IEEE International Symposium on Wireless Communications Systems (ISWCS), held in York, UK in September 2010, and the Technical Programme Chair of ISWCS 2013 in Ilmenau, Germany.