Adaptive Decision Feedback Detection with Constellation Constraints for MIMO Systems

Peng Li, *Student Member, IEEE* and Rodrigo C. de Lamare, *Senior Member, IEEE*

***Abstract*—A novel low-complexity decision feedback detection algorithm with constellation constraints (DFCC) is proposed for MIMO systems. An enhanced interference cancellation is achieved by introducing multiple constellation points as decision candidates. A complexity reduction strategy is also developed to avoid redundant processing with reliable decisions. For time-varying channels, the proposed receiver updates the filter weights using a recursive least squares (RLS) based algorithm. This efficient detector is also incorporated in a multiple branch (MB) structure to achieve a higher detection diversity order. A soft-output DFCC detector is also proposed as a component of an iterative detection and decoding receiver scheme. Simulations show that the proposed DFCC technique has a complexity as low as the adaptive DF detector while it significantly outperforms the ordered successive interference cancellation (OSIC) processing.**

***Index Terms*—spatial multiplexing, RLS algorithm, detection complexity, successive interference cancellation, V-BLAST.**

I. Introduction

Spatial multiplexed multi-input multi-output (MIMO) wireless communication has the ability to increase the channel capacity through additional data streams [1], [2], which need to be separated by a detection algorithm. The optimal maximum likelihood detection (MLD) [3] requires an exponentially increasing complexity with the system size and the modulation level. The newly developed sphere decoders (SD) [4], [5] approach the MLD performance with reduced complexity [6], however, they still have a lower bound on the complexity which is polynomial or exponential depending on the number of dimensions or the signal-to-noise ratio (SNR) [7].

Decision-driven detection [8]-[15], such as ordered successive interference cancellation (OSIC) [8]-[11] and decision feedback (DF) detectors [8], [12]-[14] are able to provide good performance with reasonable complexity. However, the application to systems with time-varying channels is still difficult due to the excessive computational load of parameter estimation. The filter and channel coefficients at the receiver need to be updated at each time instant in order to track the channel variations [13]. To address this problem, some simpler strategies have been proposed. In [13], an interpolation based channel tracking is deployed and a complexity and performance tradeoff is established. As a promising alternative, adaptive detection techniques may also be deployed for MIMO systems in time-varying channels [12]. The inter-stream interference introduced by spatial multiplexing can be suppressed by successively detecting the transmitted symbols at each time instant. A low-complexity data-aided adaptive detection technique for time-varying channels based on the generalized decision feedback equaliser GDFE [8] structure is developed in [12], the weight vectors are updated using the recursive least squares (RLS) algorithm. However, these DF algorithms have performances far from optimal and are severely limited by the effect of error propagation (EP) [18]-[21]. Recently, a multiple-branch (MB) based adaptive receiver structure [15] has been examined, this method can attain a near-optimal performance at the cost of a higher complexity than the conventional DF algorithms [8], [12].

In this paper, an innovative DF detection technique for MIMO systems is developed. With this low-complexity detector, the performance is improved through the introduction of constellation constraints (CC). In each time instance, the estimates of the symbols made by the filter are refined by a scheme that uses several selected constellation points to produce a number of tentative decisions. Thanks to this algorithm, 1) an enhanced interference cancellation can be achieved. 2) The refined symbol estimation produces better information for updating the adaptive filter weights by the RLS algorithm. 3) Unlike SDs [4], [5] which preserve several tree branches as the tentative decisions, the proposed detector preserves only one branch in the decision tree which prevents the complexity from growing exponentially. For soft-output processing, the proposed detector uses the produced tentative decisions to form a "list" [5] to compute the likelihood for each transmitted bit, the probability of decision errors is significantly reduced. Computer simulations indicate that the proposed decision feedback detection with constellation constraints (DFCC) significantly outperforms the conventional DF schemes (i.e. [12]) with very low additional detection complexity and is able to approach the optimal performance with a MB structure. The main contributions of this paper are:

- A decision feedback based algorithm is developed for detecting the signals transmitted in MIMO channels.
- A reliability checking scheme named CC is introduced to improve the performance of the detector.
- The MB processing [15] is incorporated into the scheme to achieve a higher detection diversity.
- The error performance and the detection complexity of the proposed algorithm are compared with several popular existing detection schemes.
- A soft-output DFCC is developed as a component of an iterative detection and decoding (IDD) receiver structure.

The organization of this paper is as follows. Section II gives the spatial multiplexing MIMO system model. The proposed

Peng Li and Rodrigo C. de Lamare {pl534;rcdl500}@ohm.york.ac.uk are with Department of Electronics, The University of York, Heslington, York, YO10 5DD, UK.

DFCC and its multiple branch implementation are described in section III, and is followed by a complexity comparison in section IV. The soft-output DFCC detector structure is introduced in Section V. The simulation results are given in Section VI and section VII concludes the paper.

II. Data and System Model

A spatial multiplexing MIMO system with N_T transmit antennas and N_R receive antennas is considered, where $N_R \geq N_T$. The system transmits N_T symbols which are organized into a $N_T \times 1$ vector $\boldsymbol{s}[i] = \left[s_1[i],\ s_2[i],\ \ldots,\ s_{N_T}[i]\right]^T$ at each time instant $[i]$, each transmitted signal is taken from a constellation $\mathcal{A} = \{a_1,\ a_2,\ \ldots,\ a_C\}$, where $(\cdot)^T$ denotes transpose and C denotes the number of constellation points. The output of the symbol mapper is the modulated symbol $s_k = \text{map}(\boldsymbol{b}_k)$, where $\boldsymbol{b}_k$ is a length $\log_2 C$ bit sequence. We assume that anti-Gray mapping is used. The symbol vector $\boldsymbol{s}[i]$ is then transmitted over the fading channels and the signals are collected at the receiver equipped with N_R antennas. The received signal is collected in an $N_R \times 1$ vector $\boldsymbol{r}[i] = \left[r_1[i],\ r_2[i],\ \ldots,\ r_{N_R}[i]\right]^T$ with sufficient statistics for detection

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

where the $N_R \times 1$ vector $\boldsymbol{v}[i]$ is a zero mean complex circular symmetric Gaussian noise with covariance matrix $E\left[\boldsymbol{v}[i]\boldsymbol{v}^H[i]\right] = \sigma_v^2\boldsymbol{I}$, where $E[\cdot]$ stands for expected value, $(\cdot)^H$ denotes the Hermitian operator, σ_v^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 $E\left[\boldsymbol{s}[i]\boldsymbol{s}^H[i]\right] = \sigma_s^2\boldsymbol{I}$, where σ_s^2 is the signal power. The elements h_{n_R,n_T} of the $N_R \times N_T$ channel matrix $\boldsymbol{H}$ are the complex channel gains from the n_T-th transmit antenna to the n_R-th receive antenna. The SNR per transmitted information bit is defined as

$$\frac{E_b}{N_0}\Big|_{\text{dB}} = 10\log_{10}\Big(\frac{N_R}{R\log_2 C}\cdot\frac{\sigma_s^2}{\sigma_v^2}\Big). \tag{2}$$

The total transmitted power $E_s = N_T \cdot \sigma_s^2$ which is evenly distributed to N_T transmit antennas. The N_R receive antennas collect a total power of $N_R E_s$ which carries $N_T \log_2 C$ coded bits or $RN_T \log_2 C$ information bits. $R < 1$ is the channel coding rate and $R = 1$ is assumed for the simulations without channel coding.

Nulling and Cancellation Order (NCO) of the DF

In DF detection algorithms, the optimal NCO is obtained according to the signal-to-interference-plus-noise-ratio (SINR) where the signals with higher post-detection SINR are detected first. The post-detected SINR can be computed using the linear minimum mean square error (LMMSE) detection with the following equation

$$\text{SINR}_k = \frac{\sigma_s^2|\boldsymbol{\omega}_{k,\text{MMSE}}\boldsymbol{h}_k|^2}{\sigma_s^2\sum_{l\neq k}|\boldsymbol{\omega}_{k,\text{MMSE}}\boldsymbol{h}_l|^2 + \sigma_v^2\|\boldsymbol{\omega}_{k,\text{MMSE}}\|^2}, \tag{3}$$

where $\boldsymbol{\omega}_{k,\text{MMSE}}$ is the k-th row of the MMSE matrix

$$\boldsymbol{\Omega}_{\text{MMSE}} = (\boldsymbol{H}^H\boldsymbol{H} + \sigma_v^2\boldsymbol{I})^{-1}\boldsymbol{H}^H, \tag{4}$$

and $\boldsymbol{h}_k$ is the k-th column vector of the channel matrix $\boldsymbol{H}$. In the following, this criterion is used to decide the optimal NCO.

III. Proposed Decision Feedback Detector

In the proposed DF detector, the received signal $\boldsymbol{r}[i]$ is filtered by a $N_R \times 1$ feedforward filter $\boldsymbol{\omega}_{f,k}^H[i]$ which acts as the nulling vector of the OSIC algorithm. Then for each data stream $k = 1, \ldots, N_T$, the decisions are accumulated and cancelled by the $(k-1)$-dimensional decision feedback filter $\boldsymbol{\omega}_{b,k}^H[i]$. Let $\hat{\boldsymbol{s}}[i] = \left[\hat{s}_1[i], \hat{s}_2[i], \ldots, \hat{s}_{N_T}[i]\right]^T$ represent the detected symbol vector and $u_k[i]$ denote the difference between the forward filter output and the backward filter output, which can be described as

$$u_k[i] = \boldsymbol{\omega}_{f,k}^H[i]\boldsymbol{r}[i] + \boldsymbol{\omega}_{b,k}^H[i]\hat{\boldsymbol{s}}_{k-1}[i], \tag{5}$$

where $\boldsymbol{\omega}_{b,1}^H = \boldsymbol{0}$ and the $(k-1)$-dimensional detected symbol vector is defined as $\hat{\boldsymbol{s}}_{k-1}[i] = \left[\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_{k-1}\right]^T$.

For notational convenience, the feedforward and feedback filters can be concatenated together as [12]

$$\tilde{\boldsymbol{\omega}}_k[i] = \begin{cases} \boldsymbol{\omega}_{f,k}[i], & k = 1 \\ \left[\boldsymbol{\omega}_{f,k}^T[i], \boldsymbol{\omega}_{b,k}^T[i]\right]^T, & k = 2, \ldots, N_T. \end{cases} \tag{6}$$

The input can also be concatenated as

$$\tilde{\boldsymbol{r}}_k[i] = \begin{cases} \boldsymbol{r}[i], & k = 1 \\ \left[\boldsymbol{r}^T[i], -\hat{\boldsymbol{s}}_{k-1}^T[i]\right]^T, & k = 2, \ldots, N_T. \end{cases} \tag{7}$$

Then, we can rewrite (5) as

$$u_k[i] = \tilde{\boldsymbol{\omega}}_k^H[i]\tilde{\boldsymbol{r}}_k[i]. \tag{8}$$

The weight vector for time-varying channels $\tilde{\boldsymbol{\omega}}_k^H[i]$ can be obtained by solving the standard least squares (LS) problem, the LS cost function with an exponential window is given by

$$\mathcal{J}_k[i] = \sum_{\tau=1}^{i}\lambda^{i-\tau}\Big|\hat{s}_k[\tau] - \tilde{\boldsymbol{\omega}}_k^H[i]\tilde{\boldsymbol{r}}_k[\tau]\Big|^2, \tag{9}$$

where $0 \ll \lambda < 1$ is the forgetting factor. The optimal tap weight minimizing $\mathcal{J}_k[i]$ is given by

$$\tilde{\boldsymbol{\omega}}_k[i] = \boldsymbol{\Phi}_k^{-1}[i]\boldsymbol{p}_k[i], \tag{10}$$

where the time-averaged cross correlation matrix is obtained by

$$\boldsymbol{\Phi}_k[i] = \sum_{\tau=1}^{i}\lambda^{i-\tau}\tilde{\boldsymbol{r}}_k[\tau]\tilde{\boldsymbol{r}}_k^H[\tau], \tag{11}$$

and the time-averaged cross correlation vector is defined by

$$\boldsymbol{p}_k[i] = \sum_{\tau=1}^{i}\lambda^{i-\tau}\tilde{\boldsymbol{r}}_k[\tau]\hat{s}_k^*[\tau]. \tag{12}$$

Using the RLS algorithm [22], the optimal weights in (10) can be calculated recursively. The RLS algorithm is summarised in [12] and [15].

$$\boldsymbol{q}_k[i] = \boldsymbol{\Phi}_k^{-1}[i-1]\boldsymbol{r}_k[i], \tag{13}$$

$$\boldsymbol{k}_k[i] = \frac{\lambda^{-1}\boldsymbol{q}_k[i]}{1 + \lambda^{-1}\boldsymbol{r}_k^H[i]\boldsymbol{q}_k[i]}, \tag{14}$$

$$\boldsymbol{\Phi}_k^{-1}[i] = \lambda^{-1}\boldsymbol{\Phi}_k^{-1}[i-1] - \lambda^{-1}\boldsymbol{k}_k[i]\boldsymbol{q}_k^H[i], \quad (15)$$

$$\tilde{\boldsymbol{\omega}}_k[i] = \tilde{\boldsymbol{\omega}}_k[i-1] + \boldsymbol{k}_k[i]\xi_k^*[i], \quad (16)$$

where

$$\xi_k[i] = \begin{cases} s_k[i] - \tilde{\boldsymbol{\omega}}_k^H[i-1]\tilde{\boldsymbol{r}}_k[i], & \text{Training Mode,} \\ \hat{s}_k[i] - \tilde{\boldsymbol{\omega}}_k^H[i-1]\tilde{\boldsymbol{r}}_k[i], & \text{Decision-directed Mode.} \end{cases} \quad (17)$$

This adaptive detection algorithm works in two modes. In the first mode, the filter weights are trained by the known training sequence $\boldsymbol{s}[i]$. After the filter weights converge to a certain point, the algorithm is then switched to the decision-directed mode. In this mode, the detector uses the detected symbols to update the tap weights. In this case, the quality of the detected symbols has a significant impact on the performance of this adaptive DF detector. In the following subsection the CC algorithm is introduced to obtain enhanced decisions for each stream and to allow this adaptive procedure to achieve a significant error performance improvement.

A. Constellation Constraints

The CC structure introduces a number of selected constellation points as the candidate decisions when the filter output $u_k[i]$ is determined unreliable. A selection algorithm is introduced to prevent the search space from growing exponentially. The reliability of the detected symbol is determined by the CC device. The CC device distinguishes the reliable feedback signals from the unreliable ones, which allows the DFCC to maintain the complexity at the same level of the conventional DF structure. In the following, the procedure for detecting $\hat{s}_k[i]$ for the k-th spatial stream is described, the detection of the other streams can be obtained accordingly.

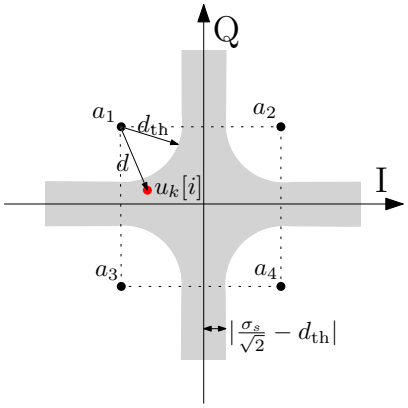

Fig. 1. The constellation constraints (CC) device. The constellation constraints procedure is invoked as the soft estimates $u_k[i]$ dropped into the shaded area.

After the system is switched to the decision-directed mode, the concatenated filter output $u_k[i]$ is checked by the CC device which is illustrated in Fig. 1. The structure is defined by the threshold distance d_{th}, which can be either a constant or a (linear) function of the signal power σ_s, and noise power σ_v. The CC device finds the constellation point which is the nearest to $u_k[i]$ according to

$$a_k[i] = \arg\min_{a_c \in \mathcal{A}} \{|u_k[i] - a_c|\}. \quad (18)$$

where a_c represents each element of the constellation. The decision is considered unreliable if one of the following equations is true

$$d > d_{\text{th}} \quad \text{when} \quad \begin{cases} \left|\text{Re}\{u_k[i]\}\right| \le \frac{\sigma_s}{\sqrt{2}} \\ \left|\text{Im}\{u_k[i]\}\right| \le \frac{\sigma_s}{\sqrt{2}} \end{cases}, \quad (19)$$

$$\begin{matrix} \left|\text{Re}\{u_k[i]\}\right| < \frac{\sigma_s}{\sqrt{2}} - d_{\text{th}} \\ \text{OR} \\ \left|\text{Im}\{u_k[i]\}\right| < \frac{\sigma_s}{\sqrt{2}} - d_{\text{th}} \end{matrix} \text{ when } \begin{cases} \left|\text{Re}\{u_k[i]\}\right| > \frac{\sigma_s}{\sqrt{2}} \\ \left|\text{Im}\{u_k[i]\}\right| > \frac{\sigma_s}{\sqrt{2}} \end{cases}, \quad (20)$$

where d denotes the distance between $u_k[i]$ and $a_k[i]$. Equation (19) represents the shaded region inside the square obtained by connecting the four a_c (the a_c are assumed to have the form, $a_c = (\pm\sigma_s/\sqrt{2} \pm j\sigma_s/\sqrt{2})$). Equation (20) denotes the shaded region outside the square. This can be further extended to higher order constellations (eg.16-QAM) where the outer tier would be very similar to (19) and (20). However, for the inner tier constellations, if $\left|a_k[i] - u_k[i]\right| \ge d_{\text{th}}$ is true, then the symbol estimate is considered unreliable. This implementation saves the computational complexity by avoiding redundant processing with reliable decisions.

Reliable: If the filter output $u_k[i]$ is dropped into the lighted area of the constellation map, the decision is considered reliable. A quantization operation $\text{Q}(\cdot)$ is then performed as

$$\hat{s}_k[i] = \text{Q}(u_k[i]). \quad (21)$$

This quantized symbol is a reliable decision for the current stream and used to compute $\xi_k[i]$ in the decision-directed mode.

Unreliable: If it is the case that $u_k[i]$ is dropped into the shaded area of the constellation map, the decision is determined unreliable. The CC processing is invoked and a candidates set is generated as

$$\mathcal{L} = \{c_1, c_2, \ldots, c_m, \ldots, c_M\} \subseteq \mathcal{A}, \quad (22)$$

The candidates are constrained by the constellation map and the vector is a selection of the M nearest constellation points to the $u_k[i]$. The size of $\mathcal{L}$ can be either fixed or adaptive with the channel condition which introduces a tradeoff between the performance and the detection complexity. The refined estimate for this unreliable decision is obtained by

$$\hat{s}_k[i] = c_{\text{opt}}, \quad (23)$$

where c_{opt} is the optimal candidate selected from $\mathcal{L}$. This refined decision will produce a more accurate $\xi_k[i]$ which minimizes the MSE in (17) in the decision-directed mode. The benefits provided by the CC algorithm are based on the assumption that the optimal feedback candidate c_{opt} is correctly selected. This selection algorithm is described as follows:

In order to find the optimal candidate, tentative decision vectors in the set $\boldsymbol{B}_k = \left\{\boldsymbol{b}_k^1, \ldots, \boldsymbol{b}_k^m, \ldots, \boldsymbol{b}_k^M\right\}$ are defined, the number of these tentative decision vectors M equals the number of selected constellation candidates. Each vector $\boldsymbol{b}_k^m$ is defined as

$$\boldsymbol{b}_k^m[i] = \left[\hat{s}_1[i], \ldots, \hat{s}_{k-1}[i], c_m, \hat{b}_{k+1}[i], \ldots, \hat{b}_{N_T}[i]\right]. \quad (24)$$

This $N_T \times 1$ vector $\boldsymbol{b}_k^m$ consists of the following parts:

- $(k-1)$-dimensional detected symbol vector $\hat{\boldsymbol{s}}_{k-1}[i]$ which is used in (7).
- A candidate symbol c_m taken from $\mathcal{L}$ for substituting the unreliable $\mathrm{Q}(u_k[i])$ in the k-th data stream.
- By concatenating the above two parts as the previous decisions, the tentative decisions of the following streams $\hat{b}_{k+1}[i], \ldots, \hat{b}_{N_T}[i]$ are subsequently obtained by the adaptive detector.

Let us define the vector with the candidate constellation point as $\check{\boldsymbol{s}}_{k,m}[i] = \left[\hat{s}_1[i], \ldots, \hat{s}_{k-1}[i], c_m\right]^T = \left[\hat{\boldsymbol{s}}_{k-1}^T[i], c_m\right]^T$. Therefore, (7) is transformed into

$$\bar{\boldsymbol{r}}_{k+1,m}[i] = \left[\boldsymbol{r}^T[i], \check{\boldsymbol{s}}_{k,m}^T[i]\right]^T, k = 1, \ldots, N_T. \tag{25}$$

The tentative decision of the $(k+1)$ stream is

$$\hat{b}_{k+1}[i] = \mathrm{Q}\Big\{\tilde{\boldsymbol{\omega}}_{k+1}^H[i]\bar{\boldsymbol{r}}_{k+1,m}[i]\Big\}. \tag{26}$$

The CC algorithm selects the best constellation point among M candidates according to the maximum likelihood rule given by

$$m_{\text{opt}} = \arg \min_{1 \le m \le M} \left\|\boldsymbol{r}[i] - \boldsymbol{H}\boldsymbol{b}_k^m[i]\right\|^2. \tag{27}$$

Then c_{opt} is chosen to replace the unreliable decision $u_k[i]$. The same filter weight $\boldsymbol{\omega}_k[i]$ is used to process all the candidates, which allows the proposed algorithm to have the computational simplicity of the adaptive DF detector. The pseudo-code of the proposed DFCC algorithm is summarized in TABLE. I.

B. Multiple-Branch Processing and Channel Estimation

The previous subsection shows that the DFCC solves (9) with optimal NCO patterns. In this subsection, a parallel branches architecture is incorporated to achieve a higher detection diversity order. The MB processing is initially introduced in [15], where a parallel receiver structure is examined by modifying the NCO of the original DF detector in an appropriate way such that the detector obtains a group of L different estimate vectors and then selects the most likely estimates. Let us employ the variable l to denote the l-th branch. It is easy to re-formulate (8) and (9) as

$$u_{k,l}[i] = \tilde{\boldsymbol{\omega}}_{k,l}^H[i]\tilde{\boldsymbol{r}}_{k,l}[i]. \tag{28}$$

$$\mathcal{J}_{k,l}[i] = \sum_{\tau=1}^{i} \lambda^{i-\tau}\left|\hat{s}_{k,l}[\tau] - \tilde{\boldsymbol{\omega}}_{k,l}^H[i]\tilde{\boldsymbol{r}'}_k[\tau]\right|^2 \tag{29}$$

At the output of the parallel branches, the decision is selected among L decision vectors as

$$\hat{\boldsymbol{s}}_{\text{opt}} = \arg \min_{1 \le l \le L} \left\|\boldsymbol{r}[i] - \boldsymbol{H}_l\hat{\boldsymbol{s}}_l[i]\right\|^2. \tag{30}$$

the Euclidean distance can be obtained directly from (27).

As we discussed above, the MIMO channel state information is required for obtaining the optimal NCO, selecting the CC candidates (27) and generating the cancellation ordering codebook for MB processing. The LS channel estimation algorithm has been investigated in [15] and [16]. For the channel estimation, if it is the case where block fading is assumed, the channel will remain constant for a number of consecutive symbols [2]. In this situation, the channel is estimated over the training sequence and then assumed constant for the remainder of the block until the next block. For time-varying channels, the channel estimation is not only used in the training sequence (training mode) but also used in the data sequence to track the channel (decision-directed mode) [12].

TABLE I
THE PSEUDO-CODE OF THE DFCC ALGORITHM

Initialization: $i = 0$

for $k = 1$ **to** N_T **do**
  $\boldsymbol{\Phi}_k^{-1}[0] = \delta^{-1}\boldsymbol{I}$; $\boldsymbol{\omega}_{f,k}[0] = \boldsymbol{1}$; $\boldsymbol{\omega}_{b,k}[0] = \boldsymbol{0}$;
end for where δ is a small positive constant.

RLS weight update: $i > 0$

$\boldsymbol{q}_k[i] = \boldsymbol{\Phi}_k^{-1}[i-1]\boldsymbol{r}_k[i]$;
$\boldsymbol{k}_k[i] = \frac{\lambda^{-1}\boldsymbol{q}_k[i]}{1+\lambda^{-1}\boldsymbol{r}_k^H[i]\boldsymbol{q}_k[i]}$;
$\boldsymbol{\Phi}_k^{-1}[i] = \lambda^{-1}\boldsymbol{\Phi}_k^{-1}[i-1] - \lambda^{-1}\boldsymbol{k}_k[i]\boldsymbol{q}_k^H[i]$;
if the system is in the training mode
  $\xi_k[i] = s_k[i] - \tilde{\boldsymbol{\omega}}_k^H[i-1]\tilde{\boldsymbol{r}}_k[i]$;
if the system is in the decision-directed mode
  $\xi_k[i] = \hat{s}_k[i] - \tilde{\boldsymbol{\omega}}_k^H[i-1]\tilde{\boldsymbol{r}}_k[i]$;
end
$\tilde{\boldsymbol{\omega}}_k[i] = \tilde{\boldsymbol{\omega}}_k[i-1] + \boldsymbol{k}_k[i]\xi_k^*[i]$;

Constellation constraints: Decision-directed mode only

for $k = 1, \ldots, N_T$
$u_k[i] = \tilde{\boldsymbol{\omega}}_k[i-1]^H\boldsymbol{r}[i]$;
if $u_k[i]$ **is unreliable**
  $\boldsymbol{\mathcal{L}} = [c_1, c_2, \ldots, c_m, \ldots, c_M]^T$;
  for $m = 1$ **to** M **do**
    $\check{\boldsymbol{s}}_{k,m}[i] = [\hat{\boldsymbol{s}}_{k-1}^T[i], c_m]^T$;
    for $q = k+1$ **to** N_T **do**
      $\bar{\boldsymbol{r}}_q[i] = \left[\boldsymbol{r}^T[i], \check{\boldsymbol{s}}_{k,m}^T[i], \hat{b}_{k+1}, \ldots, \hat{b}_{q-1}\right]^T$;
      $\hat{b}_q = \tilde{\boldsymbol{\omega}}_q^H[i]\bar{\boldsymbol{r}}_q[i]$;
    end for
  $\boldsymbol{b}_k^m[i] = \left[\check{\boldsymbol{s}}_{k,m}[i], \hat{b}_{k+1}[i], \ldots, \hat{b}_{N_T}[i]\right]$;
  end for
  $c_{\text{opt}} = \arg\min_{1 \le m \le M} \left\|\boldsymbol{r}[i] - \boldsymbol{H}\boldsymbol{b}_k^m[i]\right\|^2$;
  $\hat{s}_k[i] = c_{\text{opt}}$;
else
  $\hat{s}_k[i] = Q(u_k[i])$;
end if

IV. DETECTION COMPLEXITY

The detection complexity of the proposed DFCC and its MB version is given in this section. Essentially, we detail the complexity of the DFCC procedure with a single branch, and the complexity of the DFCC-MB can be obtained by multiply-

TABLE II
COMPUTATIONAL COMPLEXITY OF ALGORITHMS

Algorithms	**Required complex multiplications**
Non-Adaptive OSIC	$2N^3 + N^2 + N$
LMMSE -RLS	$4N^2 + 4N$
DF -RLS	$\frac{28}{3}N^2 - \frac{4}{3}$
DFCC -RLS WORST	$(\frac{28}{3}N^2 - \frac{4}{3}) + M(\frac{5}{2}N^2 - \frac{3}{2}N)$
DFCC -RLS BEST	same as "DF -RLS"
DFCC-MB -RLS	L times "DFCC -RLS"
Standard SD [6]	$\sum_{k=1}^{N} \frac{Ck\pi^{k/2}}{\Gamma(k/2+1)} d_{SD}^k + 2N^2$

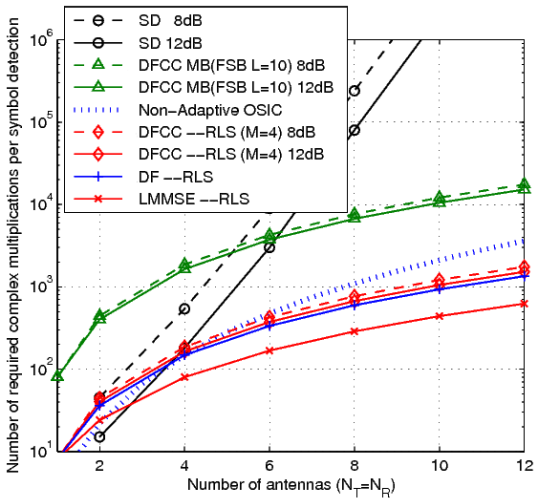

Fig. 2. Detection complexity in terms of required number of arithmetic operations per symbol detection against the number of antennas. The proposed DFCC-MB algorithm has L times the complexity of DFCC which has a similar complexity with the DF scheme.

ing the complexity of the DFCC by the number of branches L. The detailed computational complexity is shown in terms of the averaged number of required complex multiplications.

In terms of complex multiplications, the proposed algorithms and other existing schemes are represented in Table II, where $N = N_T = N_R$. The parameter M denotes the number of candidates in $\mathcal{L}$. The complexity of the SD is associated with C, the k-dimensional sphere radius d_{SD} is chosen to be a scaled version of the variance of the noise [6]. The proposed DFCC has the worst case and the best case complexities situated at the same level of the conventional DF algorithm [12]. In the worst case, all N decisions are considered unreliable, which means the $\mathcal{L}$ is generated for every stream and the CC is invoked N times which brings $M(5/2N^2 - 3/2N)$ multiplications on top of the DF algorithm. The probability of unreliable estimates is decreased as the detection diversity increases. This leads to the processing of 6.1%, 4.65%, 3.59% on average over the streams of the estimated symbol for $N = 2, 4, 8$ antennas with SNR = 16 dB, respectively. For the remaining symbols, the conventional quantization is performed. The decreased probability of invoking the CC selection suggests that the DFCC may be suitable for a MIMO system with a larger number of antenna elements.

In addition, Fig. 2 depicts the number of required complex multiplications per symbol detection. The plot compares the required operations as the number of antennas increases. Each DFCC-RLS branch has a complexity slightly above the conventional DF-RLS while it achieves a significant performance gain over the DF-RLS. The DFCC-RLS shown in the plot has the configuration of $M = 4, d_{\text{th}} = 0.5$ and the DFCC-MB-RLS has $L = 10$ branches with QPSK modulation. The low complexity of the DFCC algorithm provides an opportunity to deploy the MB structure with low computational cost, especially for systems with large antenna arrays. Compared with SD, the DFCC is simpler, this is also verified when we add more branches. In this figure, even with $L = 10$, the proposed DFCC-MB still has a complexity lower than SD when $N > 6$. Moreover, by introducing an approach named frequently selected branches (FSB) [15], it is sufficient to use a reduced number of branches. For example, we can use $L = 2$ for $N = 4$, $L = 4$ for $N = 6$ and $L = 6$ for $N = 8$ systems to obtain near-optimal performances.

V. CONSTELLATION CONSTRAINTS WITH SOFT-OUTPUT

In the following, a soft-output detector with the CCDF structure is described. Let $b_{k,j}$ be the j-th bit of the constellation symbol and $(j = 1, 2, \ldots, \log_2 C)$. We also denote $L[b_{k,j}]$ as the log-likelihood ratio (LLR) [5] value for the coded bits $b_{k,j}$. The *extrinsic* information is obtained by subtracting the dependency on $L[b_{k,j}^{(p1)}]$ as [17]

$$L[b_{k,j}^{(e1)}] = \ln \frac{\sum_{\boldsymbol{s} \in \mathcal{A}_{k,j}^1} \mathrm{P}(\boldsymbol{r}|\boldsymbol{s}) \exp\left(f(\boldsymbol{s})\right)}{\sum_{\boldsymbol{s} \in \mathcal{A}_{k,j}^0} \mathrm{P}(\boldsymbol{r}|\boldsymbol{s}) \exp\left(f(\boldsymbol{s})\right)}. \tag{31}$$

where $\mathcal{A}_{k,j}^1$ is the set of all symbol vectors that consist of bits satisfying $b_{k,j} = 1$ and $\mathcal{A}_{k,j}^0$ is similarly defined for the bits satisfying $b_{k,j} = 0$. In the complexity-reduced detection, the probability density for all the possible transmitted vectors are not available. A small set of vectors can be found by deploying the DFCC detection, the ML vector may still be found as a tentative decision. Since only a small set of symbol vectors is considered, the DFCC performance is worse than that of MAP detection when the soft-output is required. However, by appropriately selecting the tentative decisions, the DFCC performance can approach the MAP detector performance. Let $\mathcal{B}$ denote the set of tentative decisions obtained from the DFCC detector

$$\tilde{\mathcal{B}} = \boldsymbol{B}_1 \cup \boldsymbol{B}_2 \cup, \ldots, \cup \boldsymbol{B}_{N_T}, \tag{32}$$

If $L > 1$, we have

$$\mathcal{B} = \tilde{\mathcal{B}}_1 \cup \tilde{\mathcal{B}}_2 \cup, \ldots, \cup \tilde{\mathcal{B}}_L, \tag{33}$$

and the *extrinsic* information is obtained by

$$L[b_{k,j}^{(e1)}] = \ln \frac{\sum_{\boldsymbol{s}\in\mathcal{A}_{k,j}^{1}\cap\mathcal{B}} \mathrm{P}\big(\boldsymbol{r}\big|\boldsymbol{s}\big)\exp\big(f(\boldsymbol{s})\big)}{\sum_{\boldsymbol{s}\in\mathcal{A}_{k,j}^{0}\cap\mathcal{B}} \mathrm{P}\big(\boldsymbol{r}\big|\boldsymbol{s}\big)\exp\big(f(\boldsymbol{s})\big)}. \tag{34}$$

In the case that the intersection of the MAP set and the candidate vector set is empty, that is,

$$\mathcal{A}_{k,j}^{1}\cap\mathcal{B}=\emptyset \qquad \text{or} \qquad \mathcal{A}_{k,j}^{0}\cap\mathcal{B}=\emptyset, \tag{35}$$

the LLR for that specific bit can be filled with an arbitrary number with a large magnitude. Using an IDD structure [5], the *extrinsic* information $L[b_{k,j}^{(e1)}]$ is de-interleaved and fed to the channel decoder. The decoder provides the interleaved *extrinsic* information for detecting symbols in the next turbo iteration. After a number of iterations, the decision is made by slicing the LLR taken from the output of the channel decoder.

VI. Simulations

In this section, several numerical examples are used to demonstrate the overall system performance by using our algorithms. The performance is measured in terms of bit error rate (BER). In the following simulations, the transmitted vectors $\boldsymbol{s}[i]$ are grouped into blocks consisting of 500 vectors where the first $\boldsymbol{s}[1],\ldots,\boldsymbol{s}[I]$ vectors form a training sequence. We also assume that the original detection order of streams has been sorted according to the optimal order discussed in Section. II.

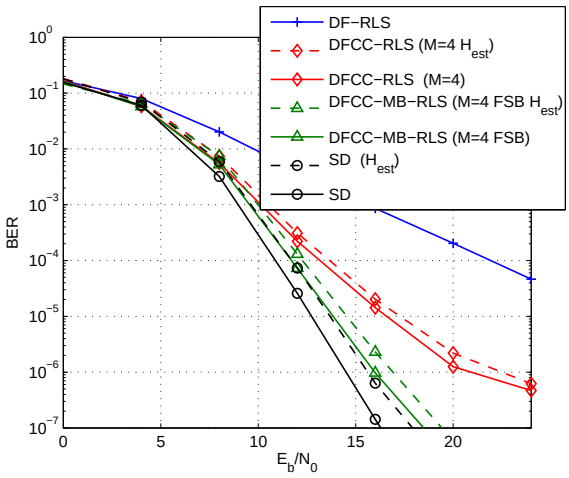

Fig. 3. BER vs. E_b/N_0. The DFCC algorithm achieves a significant performance gain compared with the DF detector. The optimal performance can be approached by employing MB in an 8×8 system configuration with perfect (solid) and imperfect channel estimation (dash). QPSK symbols, Rayleigh fading. Number of training vectors $I=50$.

The BER against E_b/N_0 plot for a system with 8×8 antennas is shown in Fig. 3. We consider that the proposed algorithms and all their counterparts operate with a channel with a block fading model. The complex channel gain matrix $\boldsymbol{H}$ with independent and identically-distributed (i.i.d) $\mathcal{CN}(0,1)$ varies from one block to another. The proposed DFCC has a 8 dB performance gain compared to the DF at the target BER 10^{-4}. The channel is assumed unknown and the LS channel estimation is applied to all the detectors indexed by H_{est}. By introducing MB (FSB $L=6$), the proposed DFCC-MB achieves a performance with about 1 dB loss from the optimal at the target BER 10^{-4}. By introducing more branches, the DFCC-MB-RLS can further improve the performance at high SNR regions.

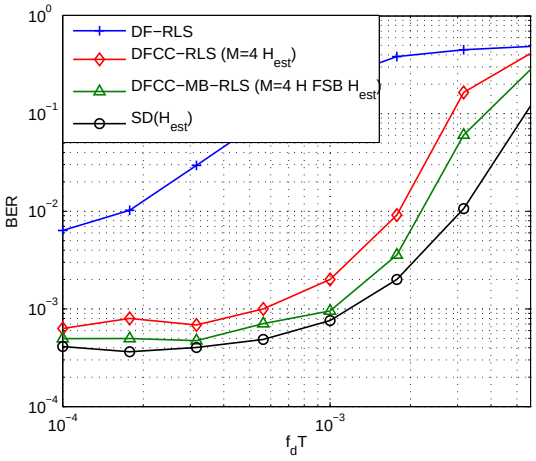

Fig. 4. Comparison of BER performance for various values of the normalized Doppler frequency f_dT when $N_T=N_R=8$ with QPSK modulation.

In order to demonstrate the tracking ability of the proposed detectors in time-varying channels, Fig. 4 presents the comparison of BER performance for various values of the normalized Doppler frequency f_dT. In this simulation, each channel between a transmit and receive antenna pair follows the Jakes' model [23]. LS channel estimation is applied to the unknown channel. Let $E_b/N_0=14$ dB for $N=4$ and $E_b/N_0=10$ dB for $N=8$. The length of the training sequence is $I=50$. The configuration of branch setting is (FSB $L=6$ [15]). The simulation results show that even with a SB, the performance of the SD with channel tracking can be approached.

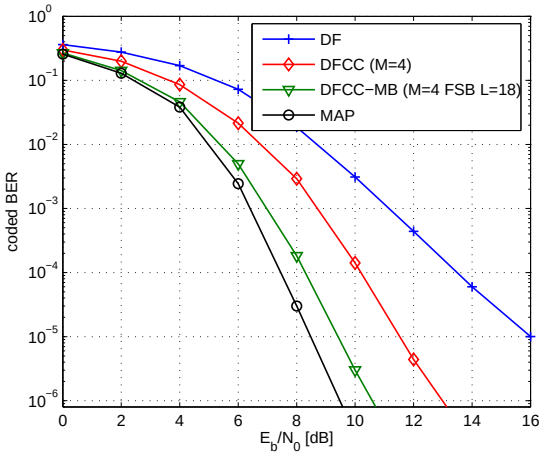

Fig. 5. Coded BER curves for QPSK over 8×8 MIMO time-varying channel with $f_dT=10^{-3}$ and the RLS algorithm. After 20 training vectors transmitted, the decision-directed mode is switched on. The interleaver size is 1000. Code rate is R = 1/2, the convolutional code has memory 2.

The curves in Fig. 5 are given for convolutionally coded BER performance with a time-varying channel. The proposed SB DFCC with $M = 4$ candidates and $d_{\text{th}} = 0.5$ improves the conventional DF detection performance by about 3.4 dB at the target coded BER of 10^{-4}. The number of branches L incorporated in the scheme introduces a tradeoff between the complexity and the performance. With (FSB L = 18) branches, the DFCC-MB detector approaches the optimal MAP detection performance with only 0.7 dB performance loss at the BER equal to 10^{-4}.

In this work, we discuss the systems with the antennas less than 10, for large MIMO and massive MIMO systems, designers can resort to reduced rank [24], [25] adaptive estimation algorithms.

VII. Conclusions

In this paper, we have developed an adaptive decision feedback-based detector for MIMO transmission systems with time-varying channels. We have presented a novel way to improve the BER performance by using the feedback with constellation constraints. This approach has the ability to effectively address the error propagation problem in decision driven interference cancellation techniques while maintaining the low complexity of adaptive detectors. The multiple branch technique can also be implemented to obtain a close to optimal performance. The soft-output requirement of the proposed detector is fulfilled by using a list of candidate decisions generated by the constellation constraint and a multiple branch procedure.

References

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585-595, Nov. 1999.

[2] G. J. Foschini, M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas", *Wireless Pers. Commun*, vol. 6, pp. 311-335, Mar. 1998.

[3] S. Verdú, *Multiuser detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[4] E. Viterbo, J. Boutros, "A universal lattice code decoder for fading channels", *IEEE Trans. Inf. Theory*, vol. 45, pp. 1639-1642, July. 1999.

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

[6] H. Vikalo, B. Hassibi, "The expected complexity of sphere decoding, part I: Theory, part II: Applications," *IEEE Trans. on Signal Process.,* 2003.

[7] J. Jalden, B. Ottersten, "On the complexity of sphere decoding in digital communications" *IEEE Trans. Signal Process.,* vol. 53, no. 4, pp. 1474- 1484, April 2005.

[8] G. Ginis, J. M. Cioffi, "On the relation between V-BLAST and the GDFE", *IEEE Commun. Lett.* 5, (9), pp. 364366, 2001.

[9] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas." *Bell Labs Tech. J.,* 1(2), 4159. 1996.

[10] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory result using V-BLAST space-time communication architecture." *Electron. Lett.*, 35(1), 1415. 1999.

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

[12] J. H. Choi, H. Y. Yu, Y. H. Lee, "Adaptive MIMO decision feedback equalization for receivers with time-varying channels", *IEEE Trans. Signal Process.*, 53, (11), pp. 42954303, 2005.

[13] Q. Sun and D. C. Cox, "Training-based channel estimation for continuous flat fading BLAST," *in proc. IEEE ICC*, Helsinki, Finland, pp. 325-329, Jun. 2002.

[14] R. C. de Lamare, R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels", *IEEE Commun. Lett.*, 7, (2), pp. 7375, 2003.

[15] R. Fa, R. C. de Lamare, "Multi-branch successive interference cancellation for MIMO spatial multiplexing systems: Design, analysis and adaptive implementation," *IET Communications*, vol.5, no.4, pp.484-494, March 2011.

[16] E. Karami, "Tracking performance of least squares MIMO channel estimation algorithm", *IEEE Trans. Commun.*, 55, (11), pp. 2201 2209, 2007.

[17] 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, March 2010.

[18] M. Chiani, "Introducing erasures in decision-feedback equalization to reduce error propagation," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 757-760, July 1997.

[19] M. Reuter, J. C. Allen, J. R. Zeidler and R. C. North, "Mitigating error propagation effects in a decision feedback equalizer", *IEEE Trans. on Commun.*, vol. 49, No.11, November 2001.

[20] R. C. de Lamare and R. Sampaio-Neto, "Minimum mean squared error iterative successive parallel arbitrated decision feedback detectors for DS-CDMA Systems,"*IEEE Trans. on Commun.*, pp. 778 - 789, May 2008.

[21] R. C. de Lamare, R. Sampaio-Neto, and A. Hjorungnes, "Joint iterative interference cancellation and parameter estimation for CDMA Systems, *IEEE Commun. Letters*, Vol. 11 No. 12, pp. 916 - 918, December 2007.

[22] S. Haykin, *Adaptive filter theory*, Third ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[23] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

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

[25] R. C. de Lamare, 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.