Fixing the Complexity of the Sphere Decoder for MIMO Detection

Luis G. Barbero, *Member, IEEE,* and John S. Thompson, *Member, IEEE*

***Abstract*—A new detection algorithm for uncoded multiple input-multiple output (MIMO) systems based on the complex version of the sphere decoder (SD) is presented in this paper. It performs a fixed number of operations during the detection process, overcoming the two main problems of the SD from an implementation point of view: its variable complexity and its sequential nature. The algorithm combines a novel channel matrix ordering with a search through a very small subset of the complete transmit constellation. A geometrically-based method is used to study the effect the proposed ordering has on the statistics of the MIMO channel. Using those results, a generalization is given for the structure this subset needs to follow in order to achieve quasi-maximum likelihood (ML) performance. Simulation results show that it has only a very small bit error rate (BER) degradation compared to the original SD while being suited for a fully-pipelined hardware implementation due to its low and fixed complexity.**

***Index Terms*—Multiple input-multiple output (MIMO), Schnorr-Euchner (SE) decoder, spatial multiplexing (SM), sphere decoder (SD), wireless communications.**

I. Introduction

THE use of multiple input-multiple output (MIMO) technology has emerged as one of the most relevant technical breakthroughs in modern wireless communications, after theoretical analysis showed that significant capacity increase could be achieved under certain conditions by using multiple antennas at both transmitter and receiver [1]. That increase in capacity can be used as a means of increasing the data rate of the system using spatial multiplexing (SM) techniques [2].

The optimum detector for those spatially multiplexed MIMO systems is the maximum likelihood detector (MLD), but its exponential complexity with the number of transmit antennas makes it unrealizable in a practical system when a large number of antennas and higher order constellations are used. In order to solve that problem, the sphere decoder (SD) has been introduced for the detection process in MIMO systems, using the underlying lattice structure of the received signal [3]. It is widely considered the most promising approach to achieve ML performance with reduced complexity compared to the MLD [4]. However, it still has an exponential lower bound in the complexity for high number of antennas and constellation orders [5]. In addition, the dependence of the actual complexity on the channel conditions and the noise level poses a problem if the SD needs to be integrated into an actual communication system, where data needs to be processed at a constant rate.

Since the introduction of the SD and its Schnorr-Euchner (SE) version [6], different alternatives have been proposed to further reduce or limit their complexity, mostly from a theoretical point of view. However, most of them still have a variable complexity. They can be classified in the following categories:

- Application of some form of lattice reduction or ordering on the channel matrix during the preprocessing stage of the SD to reduce the average complexity of the search stage [7], [8].
- Modification of the search stage of the SD to reduce the average complexity using geometric [9] or probabilistic methods [10], requiring additional operations or the calculation of limiting thresholds.
- Application of the K-Best lattice decoder [11] (equivalent to the sequential M-algorithm [12]). This approach provides a fixed complexity but it is considerably higher than the complexity of the SD in order to guarantee a quasi-ML performance [11].

In this paper, a fixed-complexity sphere decoder (FSD) based on the complex version of the SD is proposed, achieving quasi-ML performance in a fixed number of operations. The algorithm overcomes the two main problems of the SD: its variable complexity and its sequential nature. It consists of the combination of a novel channel matrix ordering and a search through a fixed small subset of the complete transmit constellation, independent of the noise level and the channel conditions. The most important part of the algorithm consists of determining the subset of the complete transmit constellation that needs to be searched. For that purpose, the effect the FSD ordering has on the channel matrix is analyzed both from a theoretical and from a simulation point of view. Using that analysis, a method is proposed to determine the subset of the transmit constellation that needs to be used for any number of antennas and constellation order.

The rest of the paper is organized as follows: Section II describes the MIMO system model and its lattice interpretation. Section III introduces the FSD algorithm and the FSD

Manuscript received June 20, 2006; revised May 11, 2007 and September 28, 2007; accepted April 7, 2008. The associate editor coordinating the review of this paper and approving it for publication was I. Collings. This paper was presented in part at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006), Toulouse, France, May 2006, and at the IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2006), Cannes, France, July 2006. This work was supported by Alpha Data Ltd. and the Scottish Funding Council through the Edinburgh Research Partnership.

L. G. Barbero was with the Institute for Digital Communications, Joint Research Institute for Signal & Image Processing, University of Edinburgh, EH9 3JL Edinburgh, UK. He is now with the ECIT, Queen's University of Belfast, BT3 9DT Belfast, UK (e-mail: l.barbero@ecit.qub.ac.uk).

J. S. Thompson is with the Institute for Digital Communications, Joint Research Institute for Signal & Image Processing, University of Edinburgh, EH9 3JL Edinburgh, UK (e-mail: john.thompson@ed.ac.uk).

Digital Object Identifier 10.1109/TWC.2008.060378.

ordering of the channel matrix. Section IV analyzes the subset of the transmit constellation that needs to be searched by the FSD. Section V discusses the performance and complexity results of the FSD for different MIMO configurations. Finally, conclusions are drawn in Section VI.

In the rest of the paper, vectors and matrices are denoted using lower-case and upper-case boldface letters, respectively, with $\mathbf{I}_N$ representing the $N \times N$ identity matrix. $(\cdot)^{\mathrm{T}}$ denotes the transpose operation and $(\cdot)^{\mathrm{H}}$ the conjugate transpose operation. The expectation operator is denoted as $\mathrm{E}[\cdot]$ and $\sim$ means *distributed as*.

II. MIMO System Model

We consider a complex-valued baseband MIMO system with M transmit and N receive antennas, denoted as $M \times N$, where $N \geq M$. Assuming symbol-synchronous sampling at the receiver and ideal timing, the N-vector of received symbols can be written, using matrix notation, as

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{v}\,,$$

where $\mathbf{s} = (s_1, s_2, \ldots, s_M)^{\mathrm{T}}$ denotes the M-vector of transmitted symbols with $\mathrm{E}[\mathbf{s}\,\mathbf{s}^{\mathrm{H}}] = (1/M)\,\mathbf{I}_M$, $\mathbf{v} = (v_1, v_2, \ldots, v_N)^{\mathrm{T}}$ is the N-vector of independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) samples $v_i \sim \mathcal{CN}(0, \sigma^2)$ with $\sigma^2 = N_0$ and $\mathbf{r} = (r_1, r_2, \ldots, r_N)^{\mathrm{T}}$ is the N-vector of received symbols. $\mathbf{H}$ denotes the $N \times M$ block Rayleigh fading channel matrix with independent elements $h_{ij} \sim \mathcal{CN}(0, 1)$ representing the complex transfer function from transmitter j to receiver i. The entries of $\mathbf{H}$ are considered to be perfectly estimated at the receiver. The transmitted symbols per antenna are taken independently from a quadrature amplitude modulation (QAM) constellation $\mathcal{O}$ of P points, representing a spatially multiplexed MIMO system. The set of all possible transmitted vectors form an M-dimensional complex constellation $\mathcal{O}^M$ of P^M vectors, which indicates the dimensionality of the system.

Since the elements of $\mathbf{H}$ are i.i.d. complex Gaussian and $\mathbf{H}$ has rank M, the set $\{\mathbf{Hs}\}$ can be considered as the complex lattice $\Lambda(\mathbf{H})$ generated by $\mathbf{H}$. The FSD proposed here is directly applied to the complex lattice so that it can be used for arbitrary complex constellations in a similar way to [13]. In addition, avoiding the more common real decomposition would result in a more efficient hardware implementation as shown for the SD in [14]. This new detector can also be applied seamlessly to the real decomposition of the system, giving a performance and complexity trade-off similar to the one presented in this paper for the complex case.

III. Fixed-Complexity Sphere Decoder (FSD)

The main concepts of the SD algorithm are briefly revised here for completeness, in order to introduce the new FSD detector. The basic idea behind the SD is to reduce the computational complexity of the MLD by searching over only those vectors of the lattice Λ that lie within a hypersphere of radius R around the received vector $\mathbf{r}$ [8]. The SD search can be represented by

$$\hat{\mathbf{s}}_{\mathrm{SD}} = \arg \min_{\substack{\mathbf{s} \in \mathcal{O}^M : \\ \|\mathbf{r} - \mathbf{H}\mathbf{s}\|^2 \leq R^2}} \|\mathbf{r} - \mathbf{H}\mathbf{s}\|^2\,, \tag{1}$$

where $\hat{\mathbf{s}}_{\mathrm{SD}} = \hat{\mathbf{s}}_{\mathrm{ML}}$, i.e. the performance of the SD is the same as that of the MLD, if, at least, one vector $\mathbf{Hs}$ is found inside the hypersphere [8].

The Euclidean distance calculation of the metric constraint in (1) can also be written, after matrix decomposition and removal of constant terms, as

$$\|\mathbf{U}(\mathbf{s} - \hat{\mathbf{s}})\|^2 \leq R^2\,, \tag{2}$$

where $\mathbf{U}$ is an $M \times M$ upper triangular matrix, with entries denoted u_{ij}, obtained through Cholesky decomposition of the Gram matrix $\mathbf{G} = \mathbf{H}^{\mathrm{H}}\mathbf{H}$ or, equivalently, QR decomposition of $\mathbf{H}$, and $\hat{\mathbf{s}} = \mathbf{H}^{\dagger}\mathbf{r}$ corresponds to the Babai point, where $\mathbf{H}^{\dagger} = (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}$ is the pseudoinverse of $\mathbf{H}$ [13].

The solution to (2) can be obtained recursively using a *depth-first* constrained tree search through M levels, starting from $i = M$ down to $i = 1$, where P child nodes, each one corresponding to one constellation point, originate from each parent node on the tree [15]. Thus, the SD traverses down the tree visiting, from each parent node, only the child nodes that satisfy

$$|s_i - z_i|^2 \leq \frac{T_i}{u_{ii}^2}\,, \tag{3}$$

where

$$z_i = \hat{s}_i - \sum_{j=i+1}^{M} \frac{u_{ij}}{u_{ii}}(s_j - \hat{s}_j)\,,$$

with $z_M = \hat{s}_M$, and

$$T_i = R^2 - \sum_{j=i+1}^{M} u_{jj}^2 |s_j - z_j|^2\,, \tag{4}$$

with $T_M = R^2$.

If a leaf node is reached (at $i = 1$) satisfying the overall metric constraint in (2), the path in the tree leading to that leaf node is stored as possible ML path and the radius is updated with its associated Euclidean distance. The SD then continues the tree search with the new metric constraint until no more paths ending in a leaf node are found inside the current hypersphere. The last path found inside the hypersphere corresponds then to the ML solution.

The reduction in complexity of the SD tree search compared to an exhaustive search comes from the fact that, at any level i, if $T_i < 0$ or no child nodes satisfy (3), the partial path down to that parent node and, consequently, all the paths originating from that node can be discarded. At that point, the SD backtracks and continues the search from a different parent node (the one closest to a leaf node, i.e. smaller i, with the largest T_i). Fig. 1 shows a simplified diagram of the tree search that would be performed in a 4×4 system with 4-QAM modulation. The curve indicates the initial metric constraint and the dashed lines indicate the paths that are discarded due to that metric constraint, reducing the total number of paths followed by the SD.

It should be noted that the order in which the child nodes originating from the same parent node are visited is relevant to the final complexity of the SD. Two possible node enumeration techniques exist, Fincke-Pohst (FP) and SE, with the latter yielding a lower complexity [8].

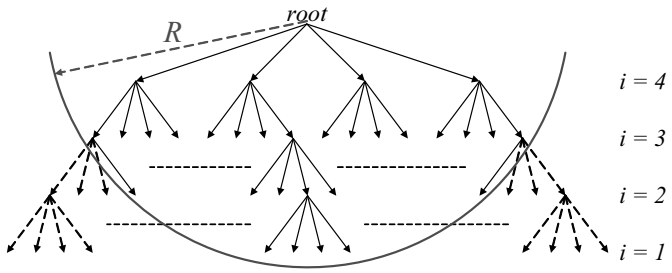

Fig. 1. Tree search diagram of the SD in a 4×4 system with 4-QAM modulation.

Although the SD provides optimal ML performance, two problems affecting its hardware implementation arise due to the *depth-first* constrained tree search:

1) The complexity of the SD depends on the number of nodes visited during the tree search. Furthermore, the number of child nodes satisfying (3) at each level i is a random variable depending on the noise level, through z_i, and on the channel conditions, through u_{ii}^2 and also indirectly through z_i. This results in the variable complexity of the SD, as shown in [8], [16], which hinders its integration into a complete communication system that would normally require data to be processed at a constant rate.
2) In addition, the existence of a metric constraint that can be updated during the tree search makes it impossible to perform parallel path searches, since future paths need to consider the results obtained from previously followed paths. This results in a sequential search where each path has to be followed to its maximum depth (depending on the metric constraint), before a search can start following any other path. This sequential nature has a dramatic effect in the detection speed (i.e. throughput) of any hardware implementation of the SD as shown in [14], [17].

A. FSD Algorithm

The new FSD proposed here overcomes the two problems of the SD by searching, independent of the noise level and the channel conditions, over only a fixed number of lattice vectors $\mathbf{Hs}$, generated by a subset of all constellation points $\mathcal{S} \subset \mathcal{O}^M$, around the received vector $\mathbf{r}$. This is equivalent to performing a fixed tree search following predetermined paths down the tree, yielding a fixed complexity algorithm. In addition, by following predefined paths down the tree, the FSD does not suffer from the sequential nature of the SD given that, potentially, all the paths can be searched in parallel.

The algorithm makes use of the statistical distribution of the random matrices involved in the SD algorithm. The channel matrix $\mathbf{H}$ has been defined as complex Gaussian, $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N \otimes \mathbf{I}_M)$, with mean $\mathrm{E}[\mathbf{H}] = \mathbf{0}$ and covariance $\mathrm{cov}[\mathbf{H}] = \mathbf{I}_N \otimes \mathbf{I}_M$. In this case, the Gram matrix $\mathbf{G} = \mathbf{H}^{\mathrm{H}}\mathbf{H}$ has a complex central Wishart distribution with N degrees of freedom, $\mathbf{G} \sim \mathcal{CW}_M(N, \mathbf{I}_M)$ [18].

The Cholesky decomposition of $\mathbf{G}$ yields an $M \times M$ upper triangular matrix $\mathbf{U}$ with independent elements such that (complex equivalent of Bartlett's decomposition) [18]:

TABLE I
$\mathrm{E}[n_i]$ FOR DIFFERENT SNRS PER BIT IN A 4×4 SYSTEM WITH 16-QAM MODULATION.

E_b/N_0 (dB)	0	10	20
$\mathrm{E}[n_4]$	3.1135	1.9098	1.1335
$\mathrm{E}[n_3]$	1.6439	1.1997	1.0049
$\mathrm{E}[n_2]$	1.1781	1.0517	1.0003
$\mathrm{E}[n_1]$	1.0	1.0	1.0

- The diagonal elements, u_{ii}, are such that $2u_{ii}^2$ are real-valued, have a Chi-square distribution with $2(N-i+1)$ degrees of freedom, $\chi^2_{2(N-i+1)}$, and $\mathrm{E}[u_{ii}^2] = N-i+1$, with $i = 1, \ldots, M$.
- The off-diagonal elements, u_{ij} with $i < j$, are independent complex Gaussian random variables $u_{ij} \sim \mathcal{CN}(0,1)$.

Therefore, the diagonal elements, u_{ii}, satisfy

$$\mathrm{E}[u_{MM}^2] < \mathrm{E}[u_{M-1M-1}^2] < \cdots < \mathrm{E}[u_{11}^2]\,, \tag{5}$$

and the metric constraints in (3), averaged over $\mathbf{H}$, $\mathbf{s}$ and $\mathbf{v}$, satisfy

$$\mathrm{E}\left[\frac{T_M}{u_{MM}^2}\right] > \mathrm{E}\left[\frac{T_{M-1}}{u_{M-1M-1}^2}\right] > \cdots > \mathrm{E}\left[\frac{T_1}{u_{11}^2}\right], \tag{6}$$

whose proof can be found in Appendix A.

It can be observed that, at each level i, (3) defines a disc on the complex plane with radius $\sqrt{T_i}/u_{ii}$ and center z_i. Therefore, the number of child nodes that could potentially be visited per parent node at each level i can be obtained by finding the number of constellation points that lie on the aforementioned disc. This can be obtained through direct calculation of the $P\,|s_i - z_i|^2$ values or decomposing the QAM constellation in concentric circles and identifying the valid points in each circle as presented in [13]. However, the number of *nodes* [1] considered per level i to find the ML solution, that we define as n_i, depends, not only on the radius $\sqrt{T_i}/u_{ii}$, but also on the reduction in R^2 that occurs every time a full path is found inside the hypersphere. That reduction is constant for all $\sqrt{T_i}/u_{ii}$ so that (6) holds. Thus, in order to get some insight into the behavior of $\mathrm{E}[n_i]$, averaging over $\mathbf{H}$, $\mathbf{s}$ and $\mathbf{v}$, we need to resort to Monte Carlo simulations.

Table I shows $\mathrm{E}[n_i]$ in a 4×4 system with 16-QAM modulation for different signal to noise ratios (SNRs) per bit $E_b/N_0 = \log_2^{-1} P \,/\, \sigma^2$. The SE version of the SD has been used with the initial radius set to infinity and updated every time a full-length path is found inside the hypersphere. It can be seen how, for all E_b/N_0, $\mathrm{E}[n_i]$ decreases with decreasing i. In addition, $n_1 = 1.0$ in all cases, since the rest of the child nodes from the same parent node at level $i = 1$ would always result in a larger Euclidean distance. The trend in Table I can also be observed in systems with different numbers of antennas and/or different constellation orders, which suggests that

$$\mathrm{E}[n_M] \geq \mathrm{E}[n_{M-1}] \geq \cdots \geq \mathrm{E}[n_1]\,. \tag{7}$$

This is due to the fact that, as the radius T_i/u_{ii}^2 of the disc defined in (3) decreases with decreasing i, less constellation

[1] In the rest of the paper, the term *nodes* is used to refer to 'child nodes visited per parent node'.

points s_i are included on that disc. This can be explained as follows: whereas in the first level, $i = M$, more *nodes* need to be visited due to the interference from the other levels, the decision-feedback equalization (DFE) performed on z_i, the increase in $\mathrm{E}[u_{ii}^2]$ and the reduction in T_i reduce the number of *nodes* in the last levels to find the ML solution [8].

Using the result in (7), the FSD assigns a fixed number of *nodes* n_i [2] per level, independent of the noise level and the channel conditions. Thus, the FSD can be uniquely identified by its distribution of *nodes* $\mathbf{n}_{\mathcal{S}} = (n_1, n_2, \ldots, n_M)^{\mathrm{T}}$. The total number of paths searched by the FSD is, therefore, $N_{\mathcal{S}} = \prod_{i=1}^{M} n_i$, where simulations in Section V show that quasi-ML performance is achieved with $N_{\mathcal{S}} \ll P^M$, i.e. $\mathcal{S}$ is a very small subset of $\mathcal{O}^M$. The n_i *nodes* on each level i are visited according to increasing distance to z_i, following the SE enumeration [6].

B. FSD Ordering of the Channel Matrix

A novel method is proposed for the ordering of the channel matrix in the preprocessing stage of the FSD. It determines the detection ordering of the signals $\hat{s}_i$ according to the distribution of *nodes*, $\mathbf{n}_{\mathcal{S}}$, that is used in the search process. The FSD ordering iteratively orders the M columns of the channel matrix $\mathbf{H}$. On the i-th iteration, considering only the signals still to be detected, the signal $\hat{s}_k$ (the index k is used to indicate that it does not necessarily coincide with the index i) with the smallest post-detection noise amplification, as calculated in [2], is selected if $n_i < P$. If $n_i = P$, the signal with the largest noise amplification is selected instead.

The steps performed in every iteration are the following (for $i = M, \ldots, 1$):

1) The matrix $\mathbf{H}_i^{\dagger} = (\mathbf{H}_i^{\mathrm{H}}\mathbf{H}_i)^{-1}\mathbf{H}_i^{\mathrm{H}}$ is calculated, where $\mathbf{H}_i = \mathbf{H}_{\mathbf{k}_{i+1}}$ is the channel matrix with the columns selected in previous iterations zeroed (represented by the index vector $\mathbf{k}_{i+1}$).
2) The signal $\hat{s}_k$ to be detected is selected according to

$$k = \begin{cases} \arg\max\limits_{j} \|(\mathbf{H}_i^{\dagger})_j\|^2, & \text{if } n_i = P \\ \arg\min\limits_{j} \|(\mathbf{H}_i^{\dagger})_j\|^2, & \text{if } n_i \neq P \ , \end{cases} \tag{8}$$

where $(\mathbf{H}_i^{\dagger})_j$ represents the j-th row of $\mathbf{H}_i^{\dagger}$ with $j \in [1, M] - \{\mathbf{k}_{i+1}\}$.

The following heuristic supports this ordering approach: if the maximum possible number of *nodes*, P, is searched in one level, the *robustness* of the signal is not relevant to the final performance, therefore, the signals that suffer the largest noise amplification can be be detected in the levels where $n_i = P$. On the other hand, in the levels where the number of *nodes* is $n_i < P$, the signals that suffer the smallest noise amplification are selected in every iteration.

IV. FSD Distribution of Nodes

The key aspect in the performance and complexity of the FSD described above is the choice of the distribution of *nodes* that form the subset $\mathcal{S}$. The distribution $\mathbf{n}_{\mathcal{S}}$ determines the level of performance that can be achieved and the reduction in complexity compared to the SD and the MLD. However, that distribution of *nodes* cannot be obtained analytically for any number of antennas and constellation sizes due to several factors. Firstly, the correlation between the values n_i, due to the DFE being applied to z_i, makes it impossible to obtain close expressions for the number of *nodes* n_i considered per level, even when no ordering is applied to the channel matrix. Secondly, the FSD ordering proposed here cannot be studied from an analytical point of view for systems with $M > 2$ due to the iterative pseudoinverse calculations. This problem has been pointed out also for the vertical-Bell Labs layered space time (V-BLAST) ordering for systems with $M > 2$ [19], [20]. In addition, the SE enumeration performed in each level affects the mathematical treatment of the problem (previous approaches to obtain an expression for the complexity of the SD have concentrated on the FP enumeration [4]).

Therefore, the aim in this section is to propose and justify a heuristic for the distribution of *nodes* $\mathbf{n}_{\mathcal{S}}$ in the FSD to achieve quasi-ML performance. In order to obtain that result, we analyze the FSD ordering using two different approaches to understand the effect it has on the post-processed (i.e. after multiplication by $\mathbf{H}^{\dagger}$) signals at the receiver and on the number of *nodes* per level in the SD.

A. Effect of the FSD Ordering on the Outage Probability

We consider a 2×2 version of the system described in Section II. In this particular case, the FSD ordering does not need to calculate the pseudoinverse of the channel matrix $\mathbf{H}^{\dagger}$. The ordering can be directly applied to the channel matrix $\mathbf{H}$, rewriting (8) as

$$k = \begin{cases} \arg\min\limits_{j} \|(\mathbf{H})^j\|^2, & \text{if } n_i = P \\ \arg\max\limits_{j} \|(\mathbf{H})^j\|^2, & \text{if } n_i \neq P \ , \end{cases}$$

where $(\mathbf{H})^j$ represents the j-th column of $\mathbf{H}$ with $j \in [1, M] - \{\mathbf{k}_{i+1}\}$ (the index i is not needed because the channel matrix is the same for the two iteration steps). In order to use the same notation as in [19], we write the channel matrix as $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2]$, where $\mathbf{h}_j = (\mathbf{H})^j$.

The received signal can be written as

$$\mathbf{r} = \mathbf{h}_1 s_1 + \mathbf{h}_2 s_2 + \mathbf{v} \, ,$$

which indicates that the FSD ordering, in the 2×2 case, can be seen as a method that selects, at the receiver, the transmitted signal s_k with the lowest overall power if $n_i = P$ and the one with the largest overall power if $n_i \neq P$ (assuming equal average power transmitted from each transmit antenna).

We consider the outage probability in terms of signal power, instead of the more common definition in terms of SNR, given that the noise power is the same at every receive antenna. In particular, we are interested in the diversity order of the post-processed signals in each step. To analyze the FSD ordering, we assume that the distribution of points used in this 2×2 system has $n_2 = P$. Thus, the signal with the lowest power is detected first and the signal with the largest power is detected second. In addition, we assume that there is no error

[2] In the rest of the paper, n_i denotes a fixed value and not a random variable.

propagation from the first detection step to the second [3]. In order to obtain the outage probability curves of the post-processed signals using the FSD ordering, we take into account that $\mathbf{h}_1$ can be written as (see [19] for details)

$$\mathbf{h}_1 = \mathbf{h}_{1\|} + \mathbf{h}_{1\perp}\,,$$

where $\mathbf{h}_{1\|}$ and $\mathbf{h}_{1\perp}$ are the components of $\mathbf{h}_1$ parallel and perpendicular to $\mathbf{h}_2$, respectively (i.e. Gram-Schmidt orthogonalization process). In addition, $\|\mathbf{h}_1\|^2 \sim \chi^2_4$ and $\|\mathbf{h}_{1\|}\|^2, \|\mathbf{h}_{1\perp}\|^2 \sim \chi^2_2$. Finally, the outage probability of $\|\mathbf{h}_1\|^2$ is written as

$$\Pr[\|\mathbf{h}_1\|^2 < x] = F_h(x) = 1 - e^{-x}(1+x)\,,$$

which is the second order maximum ratio combining (MRC) [19]. All these considerations equivalently apply to $\mathbf{h}_2$.

In addition, as it was stated in [19], the post-processed signal in the first detection step is proportional to the orthogonal component $\mathbf{h}_{k\perp}$ with k depending on the FSD ordering. Therefore, to analyze the outage probability of the signal detected in the first step, we have to take into account that the FSD ordering selects $\min[\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2]$. The signal power in that first detection step (corresponding to $i = 2$), η_2, can be written as

$$\eta_2 = \min[\|\mathbf{h}_{1\perp}\|^2, \|\mathbf{h}_{2\perp}\|^2] = (\sin^2\varphi)\min[\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2]\,,$$

where φ is the angle between $\mathbf{h}_1$ and $\mathbf{h}_2$ (Fig. 3 in [19]). The cumulative distribution function (CDF) of η_2, $F_2(x)$, can be written as

$$F_2(x) = \Pr[\eta_2 < x] = \Pr\left[\min[\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2] < \frac{x}{\sin^2\varphi}\right]. \quad (9)$$

Using concepts of order statistics [21], the distribution of $\min[\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2]$ can be expressed as $1-(1-F_h(x))^2$. Thus, (9) can be rewritten as

$$F_2(x) = \int_0^{\pi/2}\left[1 - \left\{1 - F_h\left(\frac{x}{\sin^2\varphi}\right)\right\}^2\right] f_\varphi(\varphi)d\varphi\,, \quad (10)$$

where $f_\varphi(\varphi) = \sin 2\varphi$ is the probability density function (pdf) of φ with $\varphi \in [0, \pi/2]$ [19].

Evaluating the integral in (10), the outage probability at the first detection step with FSD ordering can be written as

$$F_2(x) = 1 - \left(1 + \frac{x}{2}\right)e^{-2x}\,. \quad (11)$$

A detailed proof is given in Appendix B. Looking at the asymptotic behavior of this outage probability in the low outage probability region, we obtain

$$F_2(x) \approx \frac{3x}{2}\,,\ x \to 0\,. \quad (12)$$

Comparing this result with the asymptotic behavior of the Rayleigh distribution ($F_{\mathrm{R}}(x) \approx x$, $x \to 0$), we observe that the effect of the FSD ordering is to increase the outage probability (i.e. decrease the signal power) by 1.76 dB while keeping the same diversity order. This result is consistent with the fact that the FSD ordering, in the case under investigation, detects the signal with the lowest overall power first, therefore, causing an increase in the outage probability.

In the second detection step ($i = 1$), the signal with the largest overall power is selected and the same analysis can be used for the calculation of the outage probability. In this case, there is no need to look at the post-processed signal power. There are no additional signals to be detected which implies that the interference nulling step is not required [19]. The signal power can be written as

$$\eta_1 = \max[\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2]$$

and the outage probability can be directly expressed as

$$F_1(x) = F_h^2(x) = 1 - 2(1+x)e^{-x} + (1+x)^2e^{-2x}\,.$$

Its asymptotic behavior is

$$F_1(x) \approx \frac{x^4}{4}\,,\ x \to 0\,. \quad (13)$$

This results represents a twofold diversity increase compared to the second order MRC outage probability ($F_{\mathrm{MRC}-2}(x) \approx x^2/2$, $x \to 0$). Therefore, there is a signal power increase in the second detection step compared to the no ordering case, given the fourth order diversity in (13). This reduction in the outage probability is a direct consequence of the increase in the outage probability obtained in the first detection step. In addition, if we compare the outage probability in (13) with the fourth order MRC outage probability ($F_{\mathrm{MRC}-4}(x) \approx x^4/24$, $x \to 0$), we can observe a loss of 1.95 dB.

It should be noted that this effect is the opposite of what has been reported for the V-BLAST ordering in [19], where the ordering causes no change in diversity gain, only a shift in the outage probability curves. In particular, [19] shows that, if the original V-BLAST ordering is applied, the outage probability in the first detection step behaves asymptotically as $F_{\mathrm{V}-2}(x) \approx x/2$ with $x \to 0$, representing a reduction in the outage probability compared to the Rayleigh distribution, F_{R}. In the second detection step, the outage probability behaves asymptotically as $F_{\mathrm{V}-1}(x) \approx x^2$ with $x \to 0$, representing an increase in the outage probability compared to the second order MRC outage probability $F_{\mathrm{MRC}-2}$. Thus, comparing the results to the no ordering case, selecting the signal with the largest power in the first detection step has the effect of increasing the outage probability in the second detection step although no change in diversity occurs.

From the above results, two important conclusions can be drawn for the FSD in the case of a 2×2 system:

1) If no ordering is applied to the FSD, by looking at the outage probability curves (F_{R} and $F_{\mathrm{MRC}-2}$), it can be seen that more *nodes* would need to be considered in the first detection step (lower diversity order) compared to the second detection step (higher diversity order). That matches the result obtained in (7) looking at operation of the SD.
2) When the FSD ordering is applied, the difference in *quality* between the signals in each detection step increases. We obtain a signal power loss for the first signal but a twofold diversity increase for the second signal, as shown in (12) and (13), respectively. This indicates

[3]The assumption of no error propagation is done in order to analyze the effect the channel matrix ordering has on the signals $\hat{s}_i$ independently, as shown in [19].

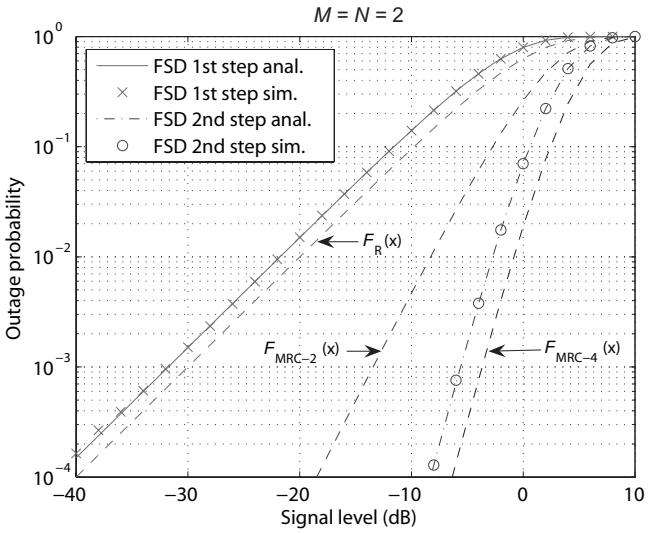

Fig. 2. Outage probability curves for the FSD ordering in a 2×2 system.

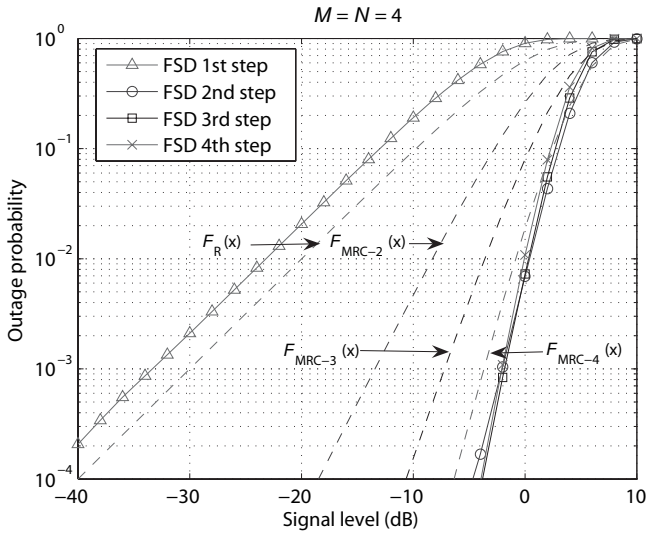

Fig. 3. Outage probability curves for the FSD ordering in a 4×4 system.

that, when the FSD is applied, more *nodes* would need to be considered in the first detection step with the advantage of reducing the number of *nodes* considered in the second detection step.

Fig. 2 shows the outage probability curves for the signals detected in each step when the FSD ordering is applied to the 2×2 case. It can be seen how the analytical results match those obtained through simulation. In the first step, the degradation of 1.76 dB can be observed compared to the Rayleigh distribution. In the second step, the diversity increase compared to the second order MRC outage probability obtained analytically is shown. In addition, the fourth order MRC outage probability has been plotted for comparison purposes [4]. Comparing it to the outage probability of the signal in the second detection step, the degradation of 1.95 dB previously calculated can be observed.

However, the same mathematical analysis cannot be done for an arbitrary number of transmit antennas and we are forced to resort to Monte Carlo simulations to determine if the behavior shown for a small system can be generalized to larger systems. That generalization is important for the FSD because it would give a heuristic justification of the distribution of *nodes* that needs to be used to obtain quasi-ML performance.

Fig. 3 shows the outage probability curves for the signals detected in each step when the FSD ordering is applied to a 4×4 system. We consider the distribution of *nodes* in the FSD to be such that $n_4 = P$ and $n_i < P$ with $i = 3, \ldots, 1$. Therefore, the signal with the lowest power is detected first while the subsequent signals are detected in descending order of signal power. The outage curves are compared with the Rayleigh distribution and the second, third and fourth order MRC outage probability (i.e. the outage curves that would be obtained in each detection step if no ordering is applied to the channel matrix). It can be seen that there is an increase in the outage probability for the signal detected in the first step, as was observed for the 2×2 case. For the remaining signals, a diversity increase (≥ 4) can be observed compared to the respective i-th order MRC diversity .

The analysis presented in this subsection suggests that, in spatially multiplexed MIMO systems, we can improve the quality of the last signals to be detected beyond the diversity order N achieved by the MLD, by detecting the signals with the lowest power in the first detection steps. This result has not been previously reported in the literature and it is of great importance to understand the distribution of *nodes* that needs to be considered in the FSD algorithm. By performing the proposed FSD ordering, we can *shift* the errors from one detection step to another, being able to predict in which step errors are more likely to happen. Thus, the FSD can achieve a quasi-ML performance with a fixed complexity, which is considerably smaller than that of the MLD.

B. Effect of the FSD Ordering on the Matrix U

In order to analyze the effect the FSD ordering has on the expected values $\mathrm{E}[u_{ii}^2]$, we initially consider the same 2×2 system. For that analysis, it should be noted that the Gram-Schmidt orthogonalization process mentioned in the previous subsection is equivalent to the Cholesky decomposition applied to $\mathbf{G}$ [23]. In particular, the (squared) diagonal elements of $\mathbf{U}$ can be written, in the general case, as

$$u_{ii}^2 = \|\mathbf{h}_i\|^2 - \sum_{k=1}^{i-1} |u_{ki}|^2 \ , \ \ i = 1, \ldots, M \, .$$

In the 2×2 case, it can be seen that, when no ordering is considered, $u_{22}^2 = \|\mathbf{h}_{2\perp}\|^2$ and $u_{11}^2 = \|\mathbf{h}_1\|^2$. When the FSD ordering is applied to $\mathbf{H}$, the (squared) diagonal elements $u_{(o)ii}^2$ can be directly written as

$$u_{(o)22}^2 = \min[\|\mathbf{h}_{1\perp}\|^2, \|\mathbf{h}_{2\perp}\|^2] = \eta_2$$

and

$$u_{(o)11}^2 = \max[\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2] = \eta_1 \, .$$

[4]The n-th order MRC outage probability can be expressed as [22]

$$F_{\mathrm{MRC-n}}(x) = 1 - e^{-x} \sum_{k=0}^{n-1} \frac{x^k}{k!} \, .$$

Therefore, the expected values $\mathrm{E}[u^2_{(o)ii}]$ can be calculated analytically using the outage probability curves from Subsection IV-A to obtain the pdfs. The results for the 2×2 case are directly given below (proof can be found in Appendix C). In the first detection step, the expected value is

$$\mathrm{E}[u^2_{(o)22}] = 5/8 = 0.625\,,$$

whereas in the second detection step is

$$\mathrm{E}[u^2_{(o)11}] = 11/4 = 2.75\ .$$

Therefore, even when the FSD ordering is applied, the expected values satisfy (5). In addition, compared to the no ordering case ($\mathrm{E}[u^2_{ii}] = N - i + 1$), $\mathrm{E}[u^2_{(o)22}] < \mathrm{E}[u^2_{22}]$ and $\mathrm{E}[u^2_{(o)11}] > \mathrm{E}[u^2_{11}]$. From the FSD algorithm point of view, the new expected values indicate the following:

1) In the first detection step, $\mathrm{E}[u^2_{(o)22}]$ indicates that the average number of constellation points that satisfies (3) is larger than in the no ordering case. Therefore, more *nodes* would generally need to be considered in the first level when the FSD ordering is used (i.e. $\mathrm{E}[n_{(o)2}] \geq \mathrm{E}[n_2]$).
2) In the second detection step, the opposite effect is found. The increase in $\mathrm{E}[u^2_{(o)11}]$ indicates that less *nodes* would need to be considered in this level.

This result for the distribution of *nodes* per level is the same that has been obtained in the previous subsection by looking at the outage probability curves.

For larger MIMO systems, the expected values $\mathrm{E}[u^2_{(o)ii}]$ can only be obtained through simulation in order to identify the evolution of the distribution of *nodes* when the FSD ordering is applied. The same trend can be observed when the number of antennas increases. That is, the FSD ordering makes the expected values of u^2_{ii} to decrease, compared to no ordering, when the signals with the lowest power are selected (i.e. levels with $n_i = P$). Simulation results have not been included due to space limitations.

C. Generalization of the Distribution of Nodes

In this subsection, a general distribution of the number of *nodes* that form the subset $\mathcal{S}$ for an arbitrary MIMO system is given in a conjecture form. The conjecture is based on the results obtained in the two previous subsections, where the effect the FSD ordering has on the system has been characterized.

Conjecture 1 (Distribution of nodes for the FSD in an arbitrary MIMO system): In an uncoded spatially multiplexed $M \times N$ system with P constellation points per transmit antenna, there exists always a distribution of *nodes* $\mathbf{n}_{\mathcal{S}}$ in the form

$$\mathbf{n}_{\mathcal{S}} = (\overbrace{1, \ldots, 1}^{l_1}, \overbrace{P, \ldots, P}^{l_P})^{\mathrm{T}} \tag{14}$$

that allows the FSD detector to achieve quasi-ML performance with the same diversity as that of the MLD and $N_{\mathcal{S}} \ll P^M$. l_P indicates the number of levels where the maximum number of *nodes* are considered and l_1 the number of levels where only one *node* is considered so that $l_1 + l_P = M$.

Justification: Although an analytical proof is currently infeasible as stated in previous sections, an analysis of the extreme distributions (when $l_1 = M$ or $l_P = M$) and of the results obtained in Subsections IV-A and IV-B are used to justify the proposed conjecture.

Firstly, in the case where $l_1 = M$, the FSD detector becomes the V-BLAST detector that belongs to the family of successive interference cancellation (SIC) detectors and has a sub-optimal BER performance. On the other hand, when $l_P = M$, the FSD detector becomes the ML detector, achieving an exact ML performance. In this case, the FSD ordering would no longer be required since the entire constellation $\mathcal{O}^M$ is searched.

In Subsections IV-A and IV-B, it has been shown how the FSD ordering reduces the number of *nodes* that need to be considered in the last levels by increasing the number of *nodes* considered in the first levels when the SD is used to obtain exact ML performance. Therefore, there exists a distribution of *nodes* $\mathbf{n}_{\mathcal{S}}$ with $l_1 \neq 0$ and $l_P < M$ achieving quasi-ML performance.

The diversity of the FSD can be justified making use of the outage probability curves obtained in Section IV-A. In particular, in the 4×4 case, Fig. 3 shows that the first signal to be detected has diversity 1 while the other signals show diversity ≥ 4. For that case, the diversity of the FSD can be justified as follows:

- The diversity of the MLD in a 4×4 system is 4. If we assume that all the *nodes* are considered only in the first level, i.e. $l_P = 1$, the degradation in performance of the FSD compared to the MLD comes from the errors in the last three levels where $n_i = 1$. Thus, the overall diversity of the FSD would be the minimum between the diversity of the MLD and the diversity of the errors in the last three levels.
- The last three levels can be seen as a 3×3 V-BLAST subsystem where the diversity of its error probability is determined by the smallest diversity of the signals to be detected due to error propagation [20]. Again, the smallest diversity of those signals has been shown to be ≥ 4 in Fig. 3.

Therefore, the diversity of the FSD is 4, equal to the diversity of the MLD, due to the increase in diversity of the signals detected in the levels where $n_i = 1$. The same justification could be used for an arbitrary MIMO system based on the outage probability curves and the choice of l_P. ■

In particular, it has been recently shown that, for $N = M$, the FSD achieves the same diversity as the MLD if the distribution of *nodes* proposed in (14) is used and $l_P \geq \sqrt{N} - 1$ [24]. This last result further shows the validity of the generalized distribution of *nodes*, presented here only as a conjecture, in order for the FSD to achieve quasi-ML performance in MIMO detection.

V. Results

The BER performance and complexity of the FSD have been obtained using Monte Carlo simulations for different constellation sizes and MIMO configurations. In all cases, the proposed FSD ordering of the channel matrix has been used. The results have been obtained using 30,000 channel realizations with 200 uncoded symbols transmitted in every channel realization.

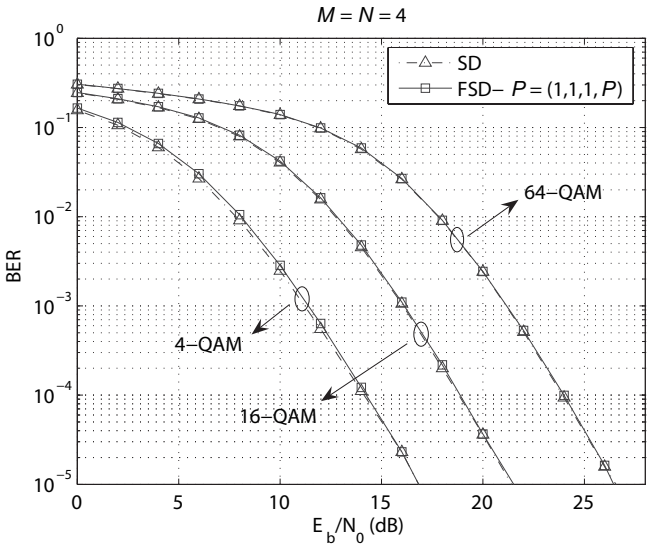

Fig. 4. BER performance of the FSD and the SD as a function of the SNR per bit in a 4×4 system.

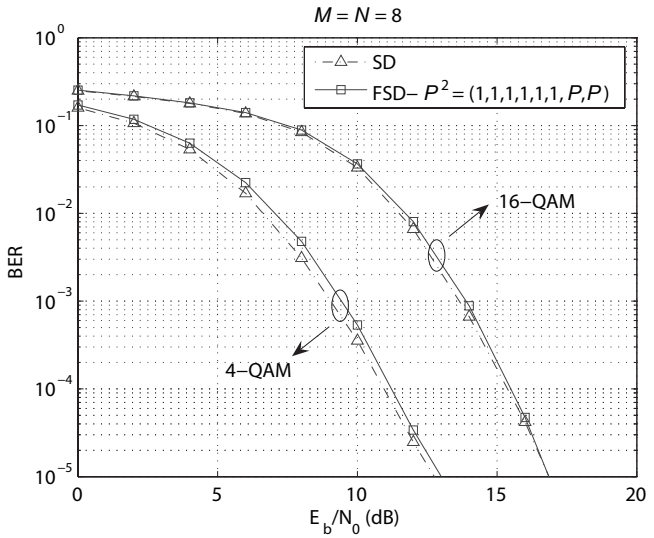

Fig. 5. BER performance of the FSD and the SD as a function of the SNR per bit in a 8×8 system.

Fig. 4 shows the BER performance of the FSD and the SD in a 4×4 system using 4-,16- and 64-QAM modulation as a function of the SNR per bit. Following the results in Section IV, the total number of paths followed by the FSD is $N_{\mathcal{S}} = P$ for a P-QAM constellation, following the distribution $\mathbf{n}_{\mathcal{S}} = (1, 1, 1, P)^{\mathrm{T}}$. It can be observed that the FSD gives practically ML performance independent of the SNR, especially for larger constellations. In particular, for 64-QAM modulation, only 64 Euclidean distances are calculated, whereas the total number of distances to be calculated by the MLD is much larger ($64^4 = 16,777,216$). The performance curves for the K-Best lattice decoder have not been included for clarity purposes, given that they show the same level of quasi-ML performance.

It should be noted that, for the specific case of $M = N = 4$, the FSD performs the same tree search as the Chase detector with list size P [25] or as a parallel detector for V-BLAST systems [26]. However, the FSD outperforms both algorithms due to the different channel matrix ordering. The algorithms in [25], [26] have been proposed as a means of improving the performance of the V-BLAST detector by performing several detections in parallel. On the other hand, the FSD represents a general algorithm proposed in the context of achieving quasi-ML performance in MIMO systems for an arbitrary constellation and for any number of antennas.

The BER performance of the FSD and the SD for a 8×8 system for 4- and 16-QAM modulation is shown in Fig. 5. In this case, the total number of paths followed by the FSD is $N_{\mathcal{S}} = P^2$ for a P-QAM constellation following the distribution $\mathbf{n}_{\mathcal{S}} = (1, 1, 1, 1, 1, 1, P, P)^{\mathrm{T}}$, i.e. $l_1 = M - 2$ and $l_P = 2$. Thus, all the possible P *nodes* are searched in the first two levels ($i = M, M-1$) and only the *node* corresponding to the closest constellation point to z_i is considered for the remaining levels. The FSD gives close to ML performance while calculating even a smaller percentage of Euclidean distances compared to the 4×4 system ($P^2/P^8 \ll P/P^4$).

For both antenna configurations, it can be observed how the performance degradation decreases when the number of constellation points per antenna P increases. This is due to the fixed structure of the FSD. At high SNR, if the ML solution is not found, the error is normally caused by wrongly taken one of the closest points to the ML point in one of the levels. Given the Gray mapping used in the QAM constellation, that causes an error in one bit. Consequently, that bit error has a greater effect on the final BER the smaller the value of $M \log_2 P$ (i.e. total number of bits per MIMO symbol).

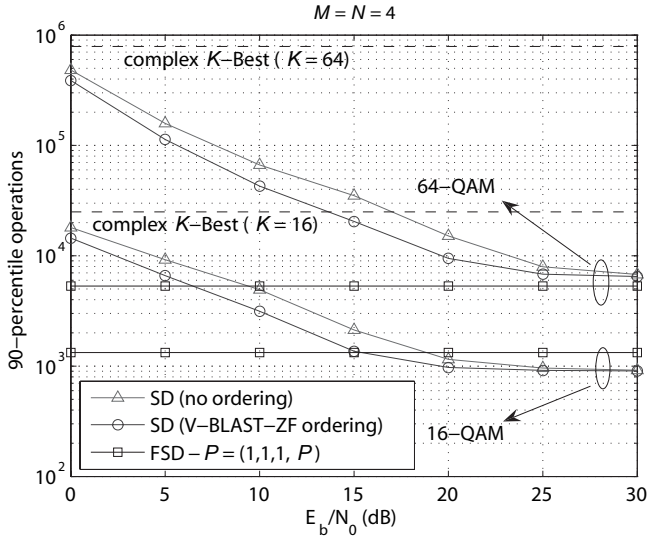

Fig. 6. Complexity of the search stage of the FSD and the SE-SD as a function of the SNR per bit in a 4×4 system.

The number of operations of the search stage of the FSD is shown in Fig. 6. The FSD is compared to the SE version of the SD with and without channel matrix ordering in a 4×4 system using 16- and 64-QAM modulation. The 90-percentile is plotted to indicate the number of operations required to perform the detection process in 90% of the cases. In order to account for the overall complexity of the algorithms, the curves include both arithmetic operations

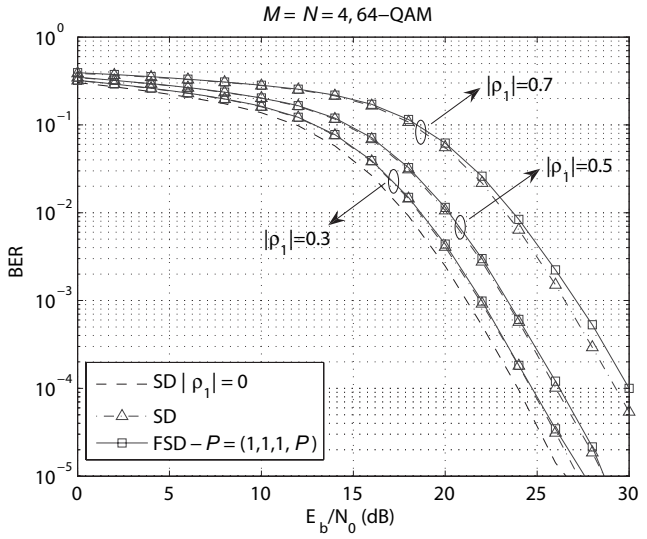

Fig. 7. BER performance of the FSD and the SD as a function of the SNR per bit in a 4×4 system in the presence of spatial correlation.

(addition, subtraction and multiplication) and logical ones (comparison, branching and sorting). The complexity of the FSD ordering stage has not been considered since it is the same as the complexity of the original V-BLAST ordering that is applied to the SD. Different optimized ordering algorithms for the latter are available in the literature that could be applied to the FSD ordering [27], [28].

Initially, the deterministic nature of the FSD can be observed, indicating its suitability for real-time hardware implementation. It can also be seen that the FSD has lower complexity than the other SDs. Only for 16-QAM and at high SNR is the number of operations of the FSD slightly higher than for the SD. The number of operations of the complex version of the K-Best lattice decoder is also plotted for comparison purposes. It can be seen how the complexity of the K-Best lattice decoder is considerably higher for both modulations.

The higher complexity of the K-Best lattice decoder is due to two factors. First, the K-Best lattice decoder, being a *breadth-first* algorithm [12], calculates the metrics of all the possible child nodes at each level increasing its computational complexity compared to that of the FSD where only some predefined child nodes have their metrics calculated. Second, the K-Best lattice decoder requires a sorting operation at each level to keep the K child nodes with the lowest metric among all the possible child nodes before proceeding down to the next level. Those sorting operations are not required in the FSD, further reducing its complexity.

Taking into account the low and fixed complexity of the FSD, a fully-pipelined design of the algorithm is possible, resulting in a considerable increase in performance in terms of detection speed as shown in [17], [29], which compare field-programmable gate array (FPGA) implementations of the SD and the FSD.

Finally, Fig. 7 shows the BER performance of the FSD and the SD in a 4×4 system with 64-QAM modulation in the presence of spatial correlation in the channel. Three different scenarios have been identified depending on the level of correlation between adjacent antennas ρ_1 (assumed to be the same at both transmitter and receiver): low correlation ($|\rho_1| \approx 0.3$), medium correlation ($|\rho_1| \approx 0.5$) and high correlation ($|\rho_1| \approx 0.7$). Details of the spatially correlated channel model and the correlation matrices for each scenario can be found in Appendix D. The distribution $\mathbf{n}_{\mathcal{S}} = (1, 1, 1, P)^{\mathrm{T}}$ has been used for the FSD. It can be seen how the presence of spatial correlation degrades the performance of both the FSD and the SD compared to the uncorrelated scenario. The effect is slightly more important on the FSD, especially in the high correlation scenario. This comes from the fact that the FSD performs the same reduced tree search independent of the channel conditions, making it more difficult to obtain the ML solution in the presence of spatial correlation, while the tree search of the SD still finds the ML solution. However, as shown in [16], the complexity of the SD increases with increasing spatial correlation in the channel, while the FSD retains the same fixed complexity. From an implementation point of view, the FSD with its fixed complexity, albeit the very small performance degradation, represents an advantage over the SD.

VI. Conclusion

A method for fixing the complexity of the SD used for MIMO detection has been presented in this paper. The proposed FSD makes it possible to achieve quasi-ML performance with a fixed complexity in systems where the MLD cannot be implemented in practice. In addition, this algorithm overcomes the two main disadvantages of the SD: its variable complexity and the sequential nature of its tree search. This results in an algorithm whose architecture can be fully-pipelined and make use of the inherent parallelism of existing hardware platforms. This makes the FSD a very suitable algorithm for hardware implementation and integration in a complete wireless system where a constant throughput needs to be guaranteed. The proposed FSD combines a novel channel matrix ordering and a search through a small subset of the transmit constellation in order to approximate ML performance. Instead of reducing the average complexity of the original SD, the focus has been on making errors more likely to occur in some specific levels, making it possible to approximate its performance with a fixed complexity algorithm.

Three directions have been identified for further research work on the FSD concept:

1) A complete analytical study of the outage probability curves of the different detection steps of the FSD for a general MIMO system would provide additional insight into the effect the FSD ordering has on the signals to be detected at the receiver. The results and methodology could then be applied to study the V-BLAST scheme or any other ordering-based MIMO detection or antenna-selection technique.
2) However, obtaining the exact outage probability curves might prove too difficult. Instead, a study of the overall diversity order of the FSD would be enough to prove its quasi-ML performance, although a completely different approach would be required. This is the main subject of recent [24] and ongoing work [30].

3) Finally, the FSD concept can also be applied to the cases where an outer code is used in the MIMO system (Turbo-MIMO systems [13]). In this case, the FSD would need to be extended to provide soft-information about the coded bits and exchange extrinsic information with the outer decoder.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers and the associate editor for their comments and suggestions that helped improve the quality and the clarity of this paper.

The authors would also like to thank Dr. Laurent Schumacher for providing the MATLAB implementation of the indoor MIMO wireless local area network (WLAN) channel model.

APPENDIX A

The expression in (6) can be obtained noting that, from the definition of the random variables T_i in (4),

$$T_M \geq T_{M-1} \geq \cdots \geq T_1 \geq 0\,, \tag{A-1}$$

where T_M is just a constant. In addition, since the elements of $\mathbf{U}$ are independent, it follows from (4) that T_i and u_{ii} are independent. Thus, the expected values in (6) can be calculated as

$$\mathrm{E}\left[\frac{T_i}{u_{ii}^2}\right] = \mathrm{E}[T_i]\,\mathrm{E}\left[\frac{1}{u_{ii}^2}\right], \tag{A-2}$$

where the random variables $1/u_{ii}^2$ have an inverse Chi-square distribution [31] satisfying

$$\mathrm{E}\left[\frac{1}{u_{MM}^2}\right] > \mathrm{E}\left[\frac{1}{u_{M-1M-1}^2}\right] > \cdots > \mathrm{E}\left[\frac{1}{u_{11}^2}\right]. \tag{A-3}$$

From here, the relationship in (6) directly follows, since the two terms on the right hand side of (A-2) decrease with decreasing i.

APPENDIX B

In order to obtain the outage probability for the signal detected in the first detection step, we first rewrite (10) as

$$\begin{aligned} F_2(x) = \quad & \int_0^{\pi/2} 2F_h\left(\tfrac{x}{\sin^2\varphi}\right)\sin 2\varphi\; d\varphi \\ & - \int_0^{\pi/2} F_h^2\left(\tfrac{x}{\sin^2\varphi}\right)\sin 2\varphi\; d\varphi\,. \end{aligned} \tag{B-1}$$

Sequentially applying the substitutions $\sin^2\varphi \to t$ and $t \to 1/t$, the integrals in (B-1) can be rewritten as

$$F_2(x) = \int_1^\infty 2\frac{F_h(xt)}{t^2}\,dt - \int_1^\infty \frac{F_h^2(xt)}{t^2}\,dt\,. \tag{B-2}$$

The solution of the first integral in (B-2) is

$$F_{2,1}(x) = 2(1 - E_2(x) - xE_1(x))\,, \tag{B-3}$$

where

$$E_k(x) = \int_1^\infty \frac{e^{-xt}}{t^k}\,dt$$

is the integral exponential function [32]. The above result can be further simplified applying the following recursive rule for the integral exponential function:

$$E_{k+1}(x) = \frac{1}{k}(e^{-x} - xE_k(x))\,,\ k = 1, 2, \ldots \tag{B-4}$$

Applying (B-4) to (B-3) to express E_2 as a function of E_1, we obtain

$$F_{2,1}(x) = 2(1 - e^{-x})\,. \tag{B-5}$$

The same method can be applied to obtain the solution of the second integral in (B-2). Firstly, we obtain

$$F_{2,2}(x) = 1 - 2E_2(x) + E_2(2x) - 2xE_1(x) + 2xE_1(2x) + \frac{x}{2}e^{-2x},$$

which can be simplified, applying (B-4), to

$$F_{2,2}(x) = 1 - 2e^{-x} + \left(1 + \frac{x}{2}\right)e^{-2x}\,. \tag{B-6}$$

Finally, combining (B-5) and (B-6), we obtain the result in (11):

$$F_2(x) = 1 - \left(1 + \frac{x}{2}\right)e^{-2x}\,. \tag{B-7}$$

APPENDIX C

The expected values $\mathrm{E}[u_{(o)ii}^2]$ are calculated using the CDFs $F_i(x)$ from (B-7) and (13).

In the first detection step, the pdf of $u_{(o)22}^2$, $f_2(x)$, is given by

$$f_2(x) = \frac{dF_2(x)}{dx} = \left(\frac{3}{2} + x\right)e^{-2x}$$

and the expected value can be expressed as

$$\mathrm{E}[u_{(o)22}^2] = \int x f_2(x)dx = \int_0^\infty \left(\frac{3x}{2} + x^2\right)e^{-2x}dx\,. \tag{C-1}$$

The integral in (C-1) can be solved applying the integration formula [32]

$$\int x^n e^{ax}dx = \frac{e^{ax}}{a^{n+1}}\left(\sum_{i=0}^{n}(-1)^i\frac{n!}{(n-i)!}(ax)^{n-i}\right),$$

with $n \in \mathbb{N}$, obtaining

$$\mathrm{E}[u_{(o)22}^2] = 5/8\,.$$

Using the same methodology, $\mathrm{E}[u_{(o)11}^2]$ can be obtained for the second detection step, using the fact that

$$f_1(x) = \frac{dF_1(x)}{dx} = 2(xe^{-x} - x(1+x)e^{-2x})\,.$$

Thus, the expected value of $u_{(o)11}^2$ is

$$\mathrm{E}[u_{(o)11}^2] = 11/4\,.$$

APPENDIX D

A simplified spatially correlated MIMO channel has been used for the case of $M = N$, modelling the channel matrix as

$$\mathbf{H}_c = (\mathbf{R})^{1/2}\mathbf{H}(\mathbf{R})^{1/2},$$

where $\mathbf{H}$ represents the uncorrelated MIMO channel as defined in Section II, $\mathbf{R}$ denotes the $M \times M$ covariance matrices representing the spatial correlation at both transmitter and receiver and $(\cdot)^{1/2}$ denotes any square root matrix such that $(\mathbf{X}^{1/2})^{\mathrm{H}}\mathbf{X}^{1/2} = \mathbf{X}$. The model assumes that the transmit and receive correlation are the same and that the antenna correlation at the receiver generated by one transmit antenna does not depend on the selected transmit antenna and is the same for all transmit antennas. The same effect is assumed for the antenna correlation generated at the transmitter. A more detailed description of the above channel model can be found in [16], [33].

With those assumptions, the $M \times M$ correlation matrix has the general form

$$\mathbf{R} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{M-1} \\ \rho_1^* & 1 & \dots & \rho_{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{M-1}^* & \rho_{M-2}^* & \dots & 1 \end{bmatrix},$$

where ρ_k represents the correlation between pairs of antennas (p, q), with $p, q = 1 \dots M$, that satisfy $p - q = k$. The conjugate value, ρ_k^*, represents the difference in phase if we consider the pairs of antennas in the opposite order (q, p).

The spatial correlation matrices have been obtained using the code available at [34], part of the IST-2000-30148 I-METRA project [35]. For the low correlation scenario, the correlation matrix is

$$\mathbf{R}_{0.3} = \begin{bmatrix} 1 & 0.24 - 0.19j & 0.11 + 0.02j & 0.05 + 0.11j \\ 0.24 + 0.19j & 1 & 0.24 - 0.19j & 0.11 + 0.02j \\ 0.11 - 0.02j & 0.24 + 0.19j & 1 & 0.24 - 0.19j \\ 0.05 - 0.11j & 0.11 - 0.02j & 0.24 + 0.19j & 1 \end{bmatrix},$$

yielding $|\rho_1| \approx 0.3$. For the medium correlation scenario, the correlation matrix is

$$\mathbf{R}_{0.5} = \begin{bmatrix} 1 & -0.50 + 0.05j & 0.21 + 0.11j & 0.01 - 0.11j \\ -0.50 - 0.05j & 1 & -0.50 + 0.05j & 0.21 + 0.11j \\ 0.21 - 0.11j & -0.50 - 0.05j & 1 & -0.50 + 0.05j \\ 0.01 + 0.11j & 0.21 - 0.11j & -0.50 - 0.05j & 1 \end{bmatrix},$$

yielding $|\rho_1| \approx 0.5$. Finally, for the high correlation scenario, the correlation matrix is

$$\mathbf{R}_{0.7} = \begin{bmatrix} 1 & 0.01 + 0.70j & -0.47 - 0.08j & 0.19 - 0.26j \\ 0.01 - 0.70j & 1 & 0.01 + 0.70j & -0.47 - 0.08j \\ -0.47 + 0.08j & 0.01 - 0.70j & 1 & 0.01 + 0.70j \\ 0.19 + 0.26j & -0.47 + 0.08j & 0.01 - 0.70j & 1 \end{bmatrix},$$

yielding $|\rho_1| \approx 0.7$.

REFERENCES

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Techn. J.*, pp. 41–59, Oct. 1996.

[2] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI International Symposium on Signals, Systems and Electronics (ISSSE '98)*, Atlanta, GA, USA, Sept. 1998, pp. 295–300.

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

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

[5] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Processing*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[6] C. P. Schnorr and M. Euchner, "Lattice basis reduction: improved practical algorithms and solving subset sum problems," *Mathematical Programming*, vol. 66, pp. 181–199, 1994.

[7] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[8] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[9] H. Artés, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: a geometrical approach to approximate ML detection," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2808–2820, Nov. 2003.

[10] W. Zhao and G. B. Giannakis, "Sphere decoding algorithms with improved radius search," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1104–1109, July 2005.

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

[12] J. B. Anderson and S. Mohan, "Sequential coding algorithms: a survey and cost analysis," *IEEE Trans. Commun.*, vol. 32, no. 2, pp. 169–176, Feb. 1984.

[13] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[14] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bölcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1566–1577, July 2005.

[15] A. D. Murugan, H. E. Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding: rediscovering the sequential decoder," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 933–953, Mar. 2006.

[16] L. G. Barbero and J. S. Thompson, "Performance of the complex sphere decoder in spatially correlated MIMO channels," *IET Commun.*, vol. 1, no. 1, pp. 122–130, Feb. 2007.

[17] ——, "Rapid prototyping of the sphere decoder for MIMO systems," in *Proc. IEE/EURASIP Conference on DSP Enabled Radio (DSPeR '05)*, vol. 1, Southampton, UK, Sept. 2005, pp. 41–47.

[18] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions*. Boca Raton, FL: Chapman & Hall / CRC, 2000.

[19] S. Loyka and F. Gagnon, "Performance analysis of the V-BLAST algorithm: an analytical approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1326–1337, July 2004.

[20] N. Prasad and M. K. Varanasi, "Analysis of decision feedback detection for MIMO rayleigh-fading channels and the optimization of power and rate allocations," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 1009–1025, June 2004.

[21] H. A. David and H. N. Nagaraja, *Order Statistics*. Hoboken, NJ: John Wiley & Sons, Inc., 2003.

[22] J. G. Proakis, *Digital Communications*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[23] G. H. Golub and C. F. V. Loan, *Matrix Computations*. London, UK: The Johns Hopkins University Press Ltd., 1996.

[24] J. Jaldén, L. G. Barbero, B. Ottersten, and J. S. Thompson, "Full diversity detection in MIMO systems with a fixed-complexity sphere decoder," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '07)*, vol. 3, Honolulu, HI, Apr. 2007, pp. 49–52.

[25] D. W. Walters and J. R. Barry, "The Chase family of detection algorithms for multiple-input multiple-output channels," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 4, Dallas, TX, Dec. 2004, pp. 2635–2639.

[26] Y. Li and Z.-Q. Luo, "Parallel detection for V-BLAST system," in *Proc. IEEE International Conference on Communications (ICC '02)*, vol. 1, New York, Apr. 2002, pp. 340–344.

[27] B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '00)*, vol. 2, Istanbul, Turkey, June 2000, pp. 737–740.

[28] H. Zhu, Z. Lei, and F. P. S. Chin, "An improved square-root algorithm for BLAST," *IEEE Signal Processing Lett.*, vol. 11, no. 9, pp. 772–775, Sept. 2004.

[29] L. G. Barbero and J. S. Thompson, "Rapid prototyping of a fixed-throughput sphere decoder for MIMO systems," in *Proc. IEEE International Conference on Communications (ICC '06)*, vol. 7, Istanbul, Turkey, June 2006, pp. 3082–3087.
[30] J. Jaldén, L. G. Barbero, B. Ottersten, and J. S. Thompson, "The error probability of the fixed-complexity sphere decoder," *in preparation*.
[31] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*. Boca Raton, FL: CRC Press, 2003.
[32] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. New York: Dover Publications, 1972.
[33] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.
[34] L. Schumacher and B. Dijkstra, "Description of a MATLAB implementation of the indoor MIMO WLAN channel model proposed by the IEEE 802.11 TGn channel model special committee," Jan. 2004. [Online]. Available: http://www.info.fundp.ac.be/~lsc/Research/IEEE_80211_HTSG_CMSC/distribution_terms.html
[35] Information Society Technologies, "IST-2000-30148 I-METRA project," http://www.ist-imetra.org.

Luis G. Barbero (M'02-S'04-M'07) was born in Barcelona, Spain, in 1976. He received the M.Sc. degree in telecommunications from the Technical University of Catalonia (UPC), Barcelona, Spain in 2000 and the Ph.D. degree in electrical engineering from the University of Edinburgh, Edinburgh, U.K. in 2006.

From 1999 to 2000, he was with Vocal Technologies Ltd., Peralada, Spain, working on the DSP implementation of modem transceivers. In 2000, he joined Massana Ltd., Dublin, Ireland, working on the physical layer testing and integration of gigabit Ethernet NIC devices. From 2001 to 2003, he was a Senior Engineer at Systemonic AG / Philips Semiconductors, Dresden, Germany, working on the design and implementation of the physical layer of a multimode WLAN system. In 2003, he joined the Institute for Digital Communications at the University of Edinburgh, Edinburgh, U.K., working towards his Ph.D. degree in the area of rapid prototyping of MIMO receive algorithms. Since 2007, he is a Research Fellow at the Institute of Electronics, Communications and Information Technology at Queen's University of Belfast, Belfast, U.K. His research interests include signal processing for wireless communications, MIMO detection algorithms, turbo-MIMO systems and implementation aspects of next-generation wireless systems.

John S. Thompson received his BEng and Ph.D. degrees from the University of Edinburgh in 1992 and 1996, respectively. From July 1995 to August 1999, he worked as a postdoctoral researcher at Edinburgh, funded by the UK Engineering and Physical Sciences Research Council (EPSRC) and Nortel Networks. Since September 1999, he has been a lecturer at the School of Engineering and Electronics at the University of Edinburgh. In October 2005, he was promoted to the position of reader. His research interests currently include signal processing algorithms for wireless systems, antenna array techniques and multihop wireless communications. He has published approximately 150 papers to date including a number of invited papers, book chapters and tutorial talks, as well as co-authoring an undergraduate textbook on digital signal processing. He is currently editor-in-chief of the IET Signal Processing journal and was recently a technical programme co-chair for the IEEE International Conference on Communications (ICC) 2007, held in Glasgow in June 2007.