# Maximum Likelihood Decoding in a Space Division Multiplexing System

Richard van Nee, Allert van Zelst and Geert Awater

Lucent Technologies Bell Labs
Zadelstede 1-10, 3431 JZ Nieuwegein, The Netherlands
Tel: +31-306097412, Fax: +31-306097498
vannee@lucent.com

**Abstract** – An analysis is made of Maximum Likelihood Decoding (MLD) in a wireless Space Division Multiplexing (SDM) link, where information is transmitted and received simultaneously over several transmit and receive antennas to achieve large data rates and high spectral efficiencies. It is proven that maximum likelihood decoding obtains a diversity order equal to the number of receive antennas, independent of the number of transmit antennas, while conventional processing techniques such as the Minimum Mean Square Error (MMSE) technique obtain a diversity order equal to the number of receive antennas minus the number of transmit antennas plus one. Hence, compared to conventional techniques, maximum likelihood decoding has a significant signal-to-noise ratio advantage which grows with the number of transmit antennas. Maximum likelihood decoding even works when the number of transmit antennas is larger than the number of receive antennas, which is not possible for conventional techniques.

## I. Introduction

The main trend in communications is more users and higher data rates per user. Hence, all new developments are aimed at increasing both the total system capacity as well as the capacity for individual users. For wireless communications with a limited amount of bandwidth, these goals require more spectrum efficient modulation techniques. In wireless local area networks, for example, we have seen the bit rates go up from 2 Mbps in the first IEEE 802.11 standard to 11 Mbps in IEEE 802.11b, using the same bandwidth [1]. The recent IEEE 802.11a standard increases the rates even further up to 54 Mbps in 20 MHz channels, using OFDM with variable QAM constellations from BPSK to 64-QAM [2]. A disadvantage of this approach to use more spectral efficient higher order modulations is that the range decreases. In addition, the links become more vulnerable to interference, which reduces the total system capacity. Hence, there is some optimum data rate per user that maximizes the system capacity. To keep increasing the rate per user *and* the system capacity, the most promising solution seems to be the use of multiple antennas.

Figure 1 shows a block diagram of a Space Division Multiplexing system, where multiple antennas are used to simultaneously transmit different data streams from one particular user, as first described in [3]. Some possible processing techniques are described in [3,4], but till now, maximum likelihood decoding has not been considered. In this paper, we derive an upperbound for the error probability of maximum likelihood decoding and show several simulation results.

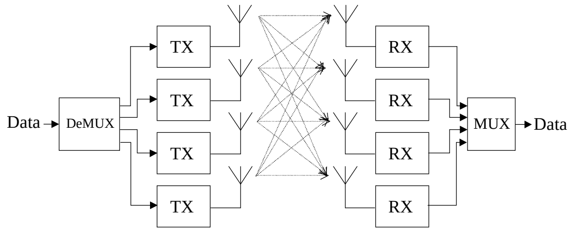



Figure 1: A Space Division Multiplexing System.

## II. Upperbound for the Symbol Error Probability

For an SDM link, the received signal vector for one particular symbol is given by:

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{H}$ is the channel matrix with $N$ rows and $M$ columns, with $N$ and $M$ being the number of receive and transmit antennas, respectively. We assume that the elements of $\mathbf{H}$ are independent zero-mean complex Gaussian variables with unit variance. $\mathbf{x}$ is the transmitted vector, consisting of $M$ QAM subsymbols. In this paper, we will refer to $\mathbf{x}$ as one symbol. The total transmitted power $P$ is equal to the sum of the powers of all elements of $\mathbf{x}$. Hence, increasing the number of transmit antennas reduces the power per antenna for a fixed total power. $\mathbf{n}$ is a vector with $N$ complex additive white Gaussian noise samples.

Assuming the receiver knows the channel $\mathbf{H}$, The maximum likelihood estimate of a symbol $\mathbf{x}$ is given by:

$$\hat{\mathbf{x}} = \arg\min_i \left\| \mathbf{r} - \mathbf{H}\mathbf{x}_i \right\| \quad (2)$$

The symbol error probability $p_s$ depends on the Euclidean distance between different received vectors $\mathbf{Hx_i}$. An upper bound on $p_s$ can be obtained by assuming all possible code words have the same minimum Euclidean distance $d_{min}$. The symbol error probability can then be written as

$$p_s \le \frac{c^M - 1}{2} erfc(\sqrt{\frac{d_{\min}^2}{4}\frac{E_s}{N_o}}) < c^M \exp(-\frac{d_{\min}^2}{4}\frac{E_s}{N_o}) \quad (3)$$

It has been assumed here that all transmit antennas use the same constellation size $c$, so the total number of code words is given by $c^M$. $d^2_{min}$ is the squared minimum Euclidean distance between two different received code words, with the average power per receive antenna normalized to one. $E_s/N_o$ is the average received signal-to-noise ratio per receive antenna.

The problem now is to find the distribution of the minimum distance. First, consider the case where two transmitted vectors differ only by a single of the $N$ $\mathbf{x}$ values. In this case, the received vector difference $\mathbf{Hx}$-$\mathbf{Hx'}$ consists of $\mathbf{H_j}(\mathbf{x_j}-\mathbf{x'_j})$ where $\mathbf{H_j}$ denotes the $j$th column of $\mathbf{H}$. The squared norm of this vector is equal to the sum of $N$ squared independent Rayleigh fading variables of $\mathbf{H_j}$, multiplied by the squared difference $|\mathbf{x_j}-\mathbf{x'_j}|^2$, which is equal to the minimum squared distance $d_c^2/M$ between two constellation points on a single receive antenna. Here, $d^2_c$ is the minimum squared distance of two constellation points with an average constellation power of one, which is listed in Table 1 for a few different modulation types. Since we assume an average received power of one per receive antenna, the average power of a single $x_j$ component is $1/M$, so the minimum squared distance for $M$ transmit antennas is $d_c^2/M$.

| Modulation | $d^2_c$ |
|---|---|
| BPSK | 4 |
| QPSK | 2 |
| 16-QAM | 4/10 |
| 64-QAM | 2/21 |

Table 1: Minimum squared distance for different modulation types with average power normalized to one.

Since the single different element between two transmitted vectors can be at any of $M$ possible locations, the minimum distance is the minimum over all $M$ columns of $\mathbf{H}$. The probability distribution of the squared minimum distance $q$ is given by:

$$P(q \le t) = 1 - \left[1 - P_{chi}(q \le t)\right]^M < MP_{chi}(q \le t) \quad (4)$$

where $P_{chi}()$ is a chi-square distribution with $2M$ degrees of freedom, which is the same distribution as is obtained in the case of classical maximal ratio combining with one transmit antenna and $M$ receive antennas. From the distribution, the probability density function can be obtained by differentiation:

$$p(q) = \frac{\partial}{\partial q}\left(1 - \left[1 - P_{chi}(q \le t)\right]^M\right) =$$
$$M\left[1 - P_{chi}(q \le t)\right]^{M-1} p_{chi}(q) \quad < \quad Mp_{chi}(q) \quad (5)$$
$$\text{because } \left[1 - P_{chi}(q \le t)\right]^{M-1} < 1$$

Hence, for the calculation of the upperbound on the bit error probability, the probability density function of $q$ can be upperbounded as $M$ times a chi-square probability density function.

If two vectors $\mathbf{x}$ and $\mathbf{x'}$ differ by more than one element, then the received vector $\mathbf{H}(\mathbf{x}-\mathbf{x'})$ consists of the sum of $K$ columns of $\mathbf{H}$, where each column of $H$ is rotated and scaled by one of the $K$ nonzero elements of $\mathbf{x}$-$\mathbf{x'}$. Assuming the worst case situation[1] that all nonzero elements of $\mathbf{x}$-$\mathbf{x'}$ have the same minimum squared Euclidean distance $d_c^2/M$, the sum of $K$ independent columns of $H$ with Gaussian distributed variables gives another set of $N$ Gaussian distributed variables with a variance that is $K$ times larger than in the case of a single different element in $\mathbf{x}$-$\mathbf{x'}$. Hence, this again gives rise to a chi-square distribution of order $2N$, but with a variance that is $K$ times larger. For small values of $q$, the probability distribution for this case can be upperbounded as $p_{chi}(q)/K$. The total number of possible locations of $K$ differences in $\mathbf{x}$-$\mathbf{x'}$ is $\binom{M}{K}$, so for all possible vector differences, an upperbound on the probability density function is:

$$p(q) < p_{chi}(q)\sum_{K=1}^{M}\binom{M}{K}\frac{1}{K} \quad (6)$$

which is nothing more than a chi-square probability density function with $2N$ degrees of freedom, multiplied by some constant which depends on the number of antennas. Now, the average symbol error probability can be obtained as:

$$p_s < \int_0^\infty c^M \exp(-\frac{d_c^2 q}{4M}\frac{E_s}{N_o})p(q)dq \quad = \quad = \frac{c^M \sum_{K=1}^{M}\binom{M}{K}\frac{1}{K}}{\left(1 + \frac{d_c^2}{4M}\frac{E_s}{N_o}\right)^N} \quad (7)$$

From the upperbound on the symbol error probability, a bit error probability upperbound can be obtained by

[1] If one or more elements have a distance larger than the minimum Euclidean distance, then the error vector still consists of $N$ Gaussian distributed elements, but with a larger variance.

assuming that for each erroneous symbol, half of the bits are in error. This leads to a bit error probability that is half of the symbol error probability.

Some interesting observations can be made from the derived upperbound: first, the diversity order is always equal to the number of receive antennas, independent of the number of transmit antennas. This is a major difference with other schemes, where the diversity order is limited to $N$-$M$+1 [5]. For relatively large signal-to-noise ratios, the symbol error probability is proportional to the inverse of the signal-to-noise ratio, raised to the power of the diversity order. Hence, maximum likelihood decoding has a significant advantage in signal-to-noise ratio over other techniques because of the larger diversity order. Notice that for a larger number of transmit antennas, the diversity order and hence the slope of the error probability curve does not change, but the curve does get multiplied by a larger constant, such that the curve shifts towards larger signal-to-noise ratios.

Figure 2 shows some bit error ratio results comparing MLD with MMSE processing. The MMSE technique estimates the transmitted vector **x** by transforming the received vector **r** to an output vector **s** as follows:

$$\mathbf{s} = (\alpha \mathbf{I} + \mathbf{H}^* \mathbf{H})^{-1} \mathbf{H}^* \mathbf{r}, \quad \alpha = \frac{M}{E_s / N_o} \qquad (8)$$

The estimate of **x** is found by slicing the individual elements of **s** to the nearest constellation points.

For MMSE, the performance only depends on the difference between the number of receive and transmit antennas, so the curves of $N$=$M$=1, 2 and 4 are identical with a diversity order of 1. For MLD, however, the diversity order grows with the number of receive antennas, thereby creating a huge signal-to-noise ratio advantage over MMSE. For $N$=$M$=4 and a bit error ratio of $10^{-3}$, for instance, MMSE requires 13 dB more signal-to-noise ratio than MLD. Notice that the theoretical upperbound for MLD is quite loose, however, it does prove the diversity order which gives the MLD curves a much steeper slope than the MMSE curves.

Notice that in all figures in this paper, we use the normalized $E_b/N_o$ per receive antenna, which is defined as:

$$\frac{E_b}{N_o} = \frac{PT_b}{N_o} \cdot N = \frac{E_s}{N_o} \cdot \frac{N}{Mk} \qquad (9)$$

where $k$ is the number of bits per subsymbol (e.g. $k$=2 for QPSK) and $T_b$ is the bit time, which is equal to the symbol time $T_s$ divided by the total amount of bits per symbol $Mk$. The $E_b/N_o$ per receive antenna is equal to the SNR of each receive antenna divided by the number of bits per symbol per receive antenna. The overall $E_b/N_o$ is equal to the $E_b/N_o$ per receive antenna divided by $N$. This means that for two systems that have the same performance in terms of $E_b/N_o$ per receive antenna, the system with the most receive antennas requires less total transmit power, so it has an advantage in the link budget. In Figure 2, for instance, all curves for MMSE are identical, but in terms of the link budget, the system with $N$=4 has a 6 dB advantage over the $N$=1 case.

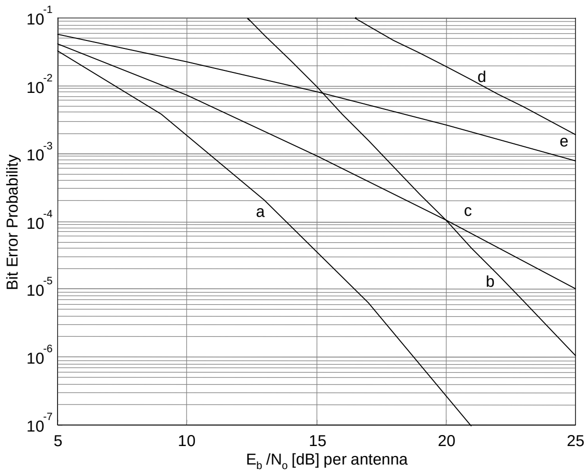
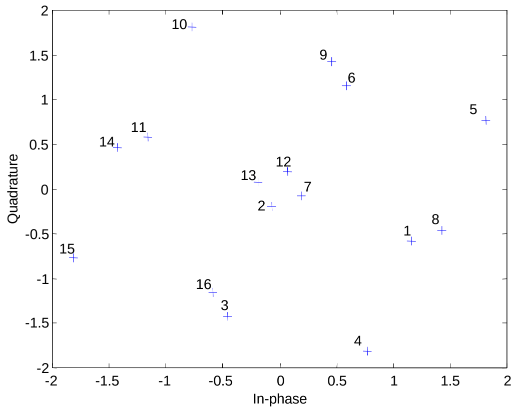
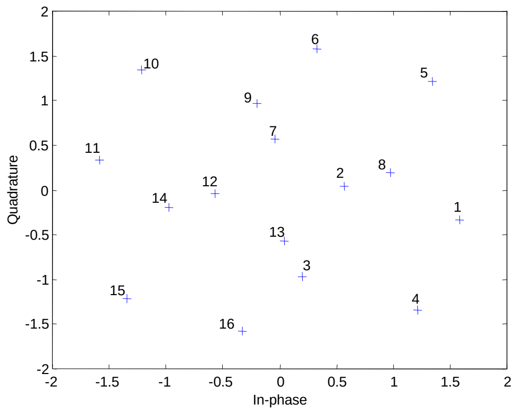



Figure 2: Bit error probability versus mean $E_b/N_o$ per receive antenna in Rayleigh fading channels for a) $N$=$M$=4, MLD simulation, b) theoretical upperbound for a, c) $N$=$M$=2, MLD simulation, d) theoretical upperbound for c, e) $N$=$M$=1, MLD and MMSE for $N$=$M$=1, 2 and 4.

A second interesting fact that follows from the derived upperbound is that maximum likelihood decoding even works when the number of transmit antennas is larger than the number of receive antennas, which is not possible for techniques which rely on the (pseudo-) inversion of (parts of) **H**. Hence, it is always possible to increase the data rate by increasing the number of transmit antennas, although at the cost of an increase in the required signal-to-noise ratio. It seems somewhat surprising that it is possible to have more transmitters than receivers, but a similar result was already found for the case of narrowband multi-user cancellation [6]. Using more transmit antennas than receive antennas can be compared with the use of higher order modulation. For instance, if 2 transmit antennas with QPSK signals are used and only one receive antenna, then the received signal in general has a 16-point constellation, with the location of the points depending on the channel matrix **H**. Figures 3 (a) and (b) show examples of these random 16-point constellations on two different receive antennas. The numbers in the plots refer to one of the 16 possible **x** inputs which led to the output shown in the figures. The figures demonstrate that one antenna is enough to recover all 4 bits of information, however, at the cost of a signal-to-noise ratio penalty, since the minimum distance of the random constellation may be quite small. When two or more receive antennas are used, the minimum distance is improved significantly, as constellation points which are spaced close together on one receive antenna may be spaced much further apart on another antenna. This is demonstrated by the example where close neighbors such as 1 and 8 in Figure 3(a) are much further apart in Figure 3(b).




(a)




(b)

Figure 3: Example of 16-point constellations on receive antennas (a) and (b) in an *M*=*N*=2 SDM system using QPSK.

Figure 4 shows the BER of BPSK with several multiple antenna combinations. Curve (a) and (b) use only one transmit antenna with one and two receive antennas, respectively, so these curves show the conventional diversity gain of increasing the number of receive antennas only. Curves (c) to (e) all have 2 receive antennas, but an increasing number of transmit antennas. Interestingly, the increased data rates only cost a slight increase in $E_b/N_o$ per antenna. For instance, there is only a 2 dB increase by going from 1 to 4 antennas (curves (b) an (e)).

The above observations pose an interesting question: for a given data rate and number of receive antennas, is it better to transmit at multiple transmit antennas or at only one antenna at the full rate, using some higher order QAM constellation? Figure 5 shows some plots for this problem with 2 receive antennas. Three systems are simulated that all transmit 4 bits per symbol, ranging from 4 BPSK symbols on 4 transmit antenas to a single 16-QAM symbol on one transmit antenna. It can be seen that 2 transmit antennas using QPSK is the best choice with an SNR advantage of about 1 and 3 dB over the BPSK and 16-QAM systems, respectively.

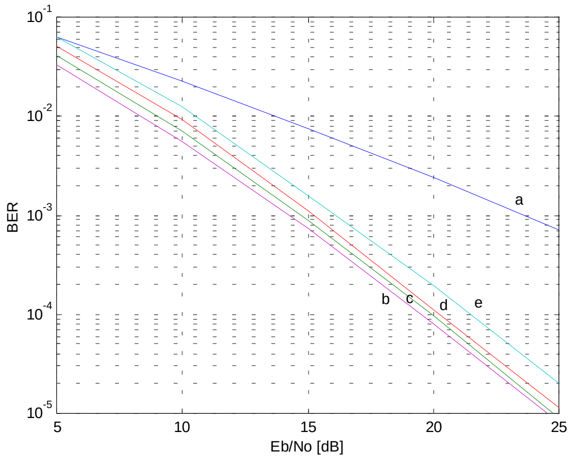



Figure 4: BER versus mean $E_b/N_o$ per receive antenna for BPSK and a) M=N=1, b) M=1, N=2, c) M=N=2, d) M=3, N=2, e) M=4, N=2.

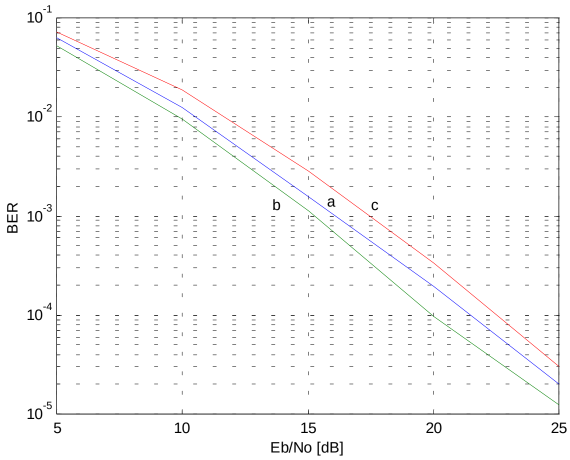



Figure 5: BER versus mean $E_b/N_o$ per receive antenna for *N*=2 and a) BPSK with *M*=4, b) QPSK with *M*=2, and c) 16-QAM with *M*=1.

Figure 6 shows another comparison between two systems with the same data rate, in this case 8 bits per symbol. Again, the system using QPSK is the best with an SNR advantage of about 1 dB over the system with 16-QAM on 2 transmit antennas.

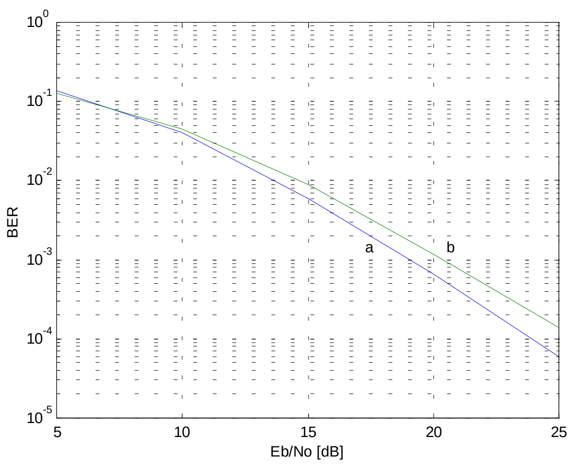



Figure 6: BER versus mean $E_b/N_o$ for N=2 and a) QPSK with M=4, b) 16-QAM with M=2.

When the number of transmit antennas is increased, the SNR penalty for higher order QAM constellations grows relatively faster than the penalty for BPSK. This is demonstreted by the curves in Figure 7 for the case of 16-QAM with 1, 2, and 3 transmit antennas. The growing SNR penalty is a direct consequence of the reduced minimum Euclidean distance because of the larger constellation sizes. For the case of 3 antennas with 16-QAM, for instance, each receive antenna sees a 4096-point constellation, while for BPSK this is only an 8-point constellation.

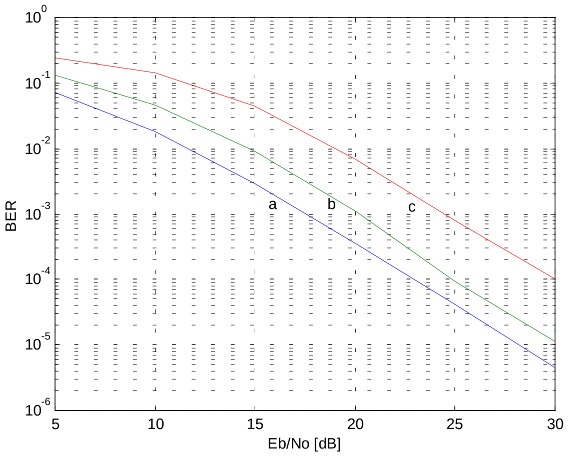



Figure 7: BER versus mean $E_b/N_o$ for 16-QAM, N=2 and a) M=1, b) M=2, and c) M=3.

## II. Complexity Issues

One of the obvious disadvantages of maximum likelihood decoding is that its complexity grows exponentially with the number of transmit antennas. However, a major advantage is that the algorithm can be made free of relatively complex multiplications by using approximations for the norm calculation. Figure 8 shows two cases where the norm is approximated as the sum of absolute values of all real and imaginary components. It can be seen that this simplification causes a degradation that is negligible for QPSK and about half a dB for 16-QAM. With this modification, the complexity of maximum likelihood decoding is in the order of $Nc^M$ complex additions, whereas the complexity of the MMSE technique, for instance, is in the order of $MN$ complex multiplications. In ASIC or FPGA implementations, an 8-bit complex multiplier is about 16 times more complex than a complex adder, so the $MN$ complex multiplications are equivalent in complexity to $16MN$ complex additions. Using this comparison rule, the implementation complexity of MLD is actually less than that of MMSE for a system with $M$=$N$=4 and QPSK ($c$=4). For larger number of antennas and/or constellation sizes, MLD is relatively more complex than MMSE. Further simplifications exist, however, which make it possible to reduce the complexity at the cost of a reduced performance [7].

## II. Conclusions

We proved that maximum likelihood decoding of an SDM system achieves a diversity factor equal to the number of receive antennas, independent of the number of transmit antennas. The reasonable implementation complexity combined with the superior performance makes maximum likelihood decoding a highly attractive technique for SDM systems, especially when combined with OFDM to gain delay spread robustness, as described in [8].

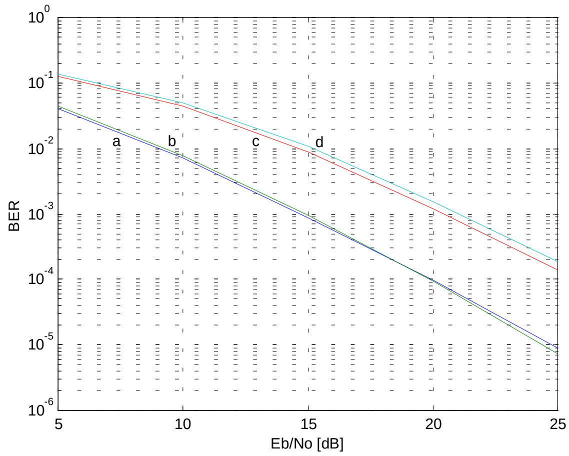



Figure 8: BER versus mean $E_b/N_o$ for M=N=2, a) QPSK exact, b) QPSK with norm approximation, c) 16-QAM exact, d) 16-QAM with norm approximation.

**REFERENCES**


[1] R. van Nee, G. Awater, M. Morikura, H. Takanashi, M. Webster, and K. Halford, “New High Rate Wireless LAN Standards,” *IEEE Communications Magazine*, Vol. 37, No. 12, pp. 82–88, Dec. 1999.

[2] R. van Nee and R. Prasad, *OFDM for Mobile Multimedia Communications*, Boston: Artech House, Dec. 1999.

[3] G. J. Foschini, “Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multi-Element Arrays,” *Bell Labs Technical Journal*, Vol. 1, No. 2, Autumn 1996, pp. 41–59.

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

[5] J. H. Winters, J. Salz and R. D. Gitlin, “The Impact of Antenna Diversity on the Capacity of Wireless Communication Systems,” *IEEE Trans. on Communications*, Vol. 42, No. 2, Feb./Mar./Apr. 1994, pp. 1740–1751.

[6] G. J. M. Janssen, “Robust Receiver Techniques for Interference-Limited Radio Channels,” *PhD Thesis*, ISBN 90-407-1676-5, Delft University Press, Delft, The Netherlands, 1998.

[7] G. Awater, A. van Zelst, and R. van Nee, “Reduced Complexity Space Division Multiplexing Receivers,” *IEEE VTC 2000*, Tokyo, Japan, May 2000.

[8] A. van Zelst, R. van Nee, and G. Awater, “Space Division Multiplexing for OFDM Systems,” *IEEE VTC 2000*, Tokyo, Japan, May 2000.