Reduced-Complexity Transmit/Receive-Diversity Systems

Andreas F. Molisch, Moe Z. Win and Jack H. Winters
AT&T Laboratories - Research
200 Laurel Av., Middletown, NJ, USA
email: Andreas.Molisch@tuwien.ac.at
A. F. Molisch was on leave from INTHFT, Technische Universitaet Wien

Abstract

We consider wireless systems with transmit and receive diversity. For reduction of complexity, we propose to use hybrid selection/maximal ratio transmission at one link end, choosing L out of N antennas. We analyze the performance of such systems, giving analytical bounds and comparing them to computer simulations. Outage probability, symbol error probability, and capacity are shown. We demonstrate that in typical cases, a small number of used antennas L is sufficient to achieve considerable performance gains.

1 Introduction

Transmit diversity schemes were first proposed in [20], [19] for the enhancement of transmission quality in mobile radio systems. References [18], [4] demonstrated how transmit diversity could be combined with receive diversity to achieve enormous capacity gains. However, a full exploitation of these gains requires the use of space-time-codes [14], which precludes their use in existing systems, where backward compatibility must be assured.

An alternative way for exploiting multiple-input - multiple-output (MIMO) systems is the use of transmit and receive diversity purely for link-quality improvement, exploiting the diversity effect. In such a system, the signals supplied to the different transmit antennas are weighted replicas of a single bit stream (which might be coded or uncoded). The ideal weights can be determined by matching them to the channel, resulting in maximal ratio transmission (MRT) [7] Similarly, at the receiver, "standard" maximal ratio combining is employed, using linear combinations of the signals obtained at the different receive antennas. It has been shown that with N_t transmit and N_r receive antennas, a diversity degree of $N_t N_r$ can be achieved [3]. Since it employs no special type of coding, any standard (single-antenna) receiver can detect the transmitted signal (albeit with a smaller diversity degree and thus reduced quality).

The main disadvantage of MRT (MRC) is the fact that it requires N_t (N_r) complete RF chains. There are numerous situations where this high degree of hardware expense is undesirable. On the other hand, simple selection diversity usually exhibits worse performance than MRC. This has led to the introduction of hybrid selection / maximum ratio combining H-S/MRC [6], [16], [17], [2]. In this scheme, the best L out of N antennas are selected, and then combined. Thus, only L RF chains are required. Up to now, this scheme has been considered only for the receiver case.

In this paper, we are treating a transmit/receive diversity system where the transmitter uses hybrid selection/maximal ratio transmission (H-S/MRT), while the receiver uses MRC. We will analyze the performance of such a system in terms of signal-to-noise ratio (SNR), symbol error probability (SEP), and capacity.

2 System and channel model

Figure 1 shows the generic system that we are considering. A bit stream is sent through an encoder, and a modulator. A multiplexer switches the modulated signals to the best L_t out of N_t available antenna branches. For each selected branch, the signal is multiplied by a complex coefficient u whose actual value depends on the current channel realization. In a real system, the signals are subsequently upconverted to passband, amplified by a power amplifier, and filtered. For our model, we omit these stages, as well as their equivalents at the receiver, and treat the whole problem in equivalent baseband. Note, however, that it exactly these parts that are most expensive and make the use of reduced-complexity systems desirable.

Next, the signal is sent over a quasi-static flat-fading channel. We also assume that the fading at all antenna elements is independent; its statistics obey a Rayleigh distribution. We denote the $N_t \times N_r$ matrix of the channel as

$$H = \begin{pmatrix} h_{11} & h_{12} & . & . & . & h_{1N_r} \\ h_{21} & h_{22} & . & . & . & h_{2N_r} \\ . & . & . & . & & . \\ h_{N_t1} & h_{N_t2} & & & & h_{N_tN_r} \end{pmatrix}. \quad (1)$$
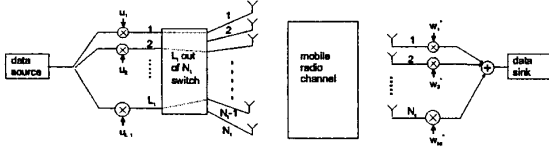
Figure 1. System model

If the channel is Rayleigh fading, the h_{ij} are independent identically distributed zero-mean, circularly symmetric complex Gaussian random variables with unit variance, i.e. the real and imaginary part each have variance $1/2$. Consequently, the power carried by each transmission channel (h_{ij}) is chi-square distributed with 2 degrees of freedom.

The output of the channel is polluted by additive white Gaussian noise, which is assumed to be independent at all receiver antenna elements. The received signals are multiplied by complex weights w^* at all antenna elements (where superscript $*$ denotes complex conjugation), and combined before passing a decoder/detector. We assume that both transmitter and receiver have perfect knowledge of the channel. This is, of course, an idealization that can only be approximated even in slowly fading channels.

We note at this point that some additional simplifying assumptions are made. They are standard assumptions in the analysis of multiple antenna systems, but still their validity must be checked when applying the above theoretical results to practical systems: (i) The fading at the different antenna elements was assumed to be independent, identically distributed Rayleigh fading or Nakagami fading with integer m-parameter. (ii) The fading is assumed to be frequency flat.

3 Computation of performance

3.1 Channel statistics and optimum weights

We first have to determine the optimum antenna weights, and the statistics of the fading channel. The easiest way for deriving the optimum weights is a singular value decomposition of the channel matrix H [5] $H = U\Lambda W^*$, where Λ is a diagonal matrix containing the singular values, and U and W^* are unitary matrices containing the left and right singular vectors, respectively. The optimum transmit weight vector $\vec{u}$ and optimum receive weight vector $\vec{w}^*$, respectively, can now be shown to be the left and right singular vectors belonging to the largest singular value [3]. The effective SNR is given by the square of this singular value, i.e. the eigenvalue of $HH^\dagger$, where superscript $\dagger$ denotes Hermitian transpose. Note that this derivation assumes the use of all available antennas at both transmitters and receivers.

Our goal here is to determine the performance when only a subset of the antennas are used. For this, we have to define a set of matrices $\widetilde{H}$, where $\widetilde{H}$ is created by striking $N_t - L_t$ rows from H, and $S(\widetilde{H})$ denotes the set of all possible $\widetilde{H}$, whose cardinality is $\binom{N_t}{L_t}$. The achievable SNR of the reduced-complexity system is now

$$SNR_{\text{H-S/MRC}} = \max_{S(\widetilde{H})}\left(\max_i(\widetilde{\lambda}_i^2)\right) \qquad (2)$$

where the $\widetilde{\lambda}_i$ are the singular values of $\widetilde{H}$.

An analytical solution $SNR_{\text{H-S/MRC}}$ does not seem to be easily obtainable. However, we can derive upper and lower bounds. We start out by stating that

$$\frac{1}{\min(L_t, N_r)}\sum_i \widetilde{\lambda}_i^2 \le \max_i(\widetilde{\lambda}_i^2) \le \sum_i \widetilde{\lambda}_i^2 \qquad (3)$$

i.e., the achievable SNR for a certain modified channel matrix $\widetilde{H}$ is lower-bounded by the *average* of the nonzero eigenvalues, and upper-bounded by the *sum* of the nonzero eigenvalues of $\widetilde{H}$. We thus can bound the SNR of the selective-transmit - receive diversity system by finding

$$\gamma_{\text{bound}} = \max_{S(\widetilde{H})}\left(\sum_i \widetilde{\lambda}_i^2\right) = \max_{S(\widetilde{H})}\left(\sum_i\sum_j |\widetilde{h}_{ij}|^2\right) \qquad (4)$$

We note here also that the antenna combination that gives the maximum $\sum_i \widetilde{\lambda}_i^2$ is not necessarily the antenna combination that gives the maximum $\max_i(\widetilde{\lambda}_i^2)$.[1] However, the bounds of Eq. 3 remain valid when the maximization over all antenna combinations is applied to them.

Now the maximization can also be interpreted as being performed over various combinations of L_t out of N_t rows, while the columns of the matrix always have dimension N_r. Thus, $\gamma_i = \sum_{j=1}^{N_r} |\widetilde{h}_{ij}|^2$ are (henceforth normalized) chi-square distributed random variables with N_r degrees of freedom. Note that the γ_i can be interpreted as the SNRs that are achieved when only the i-th antenna is transmitting, and the receiver uses MRC. The joint statistics of the *ordered* SNRs $\gamma_{(i)}$ can be shown to be [17]

$$p_{\gamma_{(i)}}(\gamma_{(1)}, \gamma_{(2)}, \dots (N_t)) = N_t! \prod_{i=1}^{N_t} \frac{1}{\Gamma(N_r)}\gamma_{(i)}^{N_r-1}e^{-\gamma_{(i)}} \qquad (5)$$

for $\gamma_{(1)} > \gamma_{(2)} > \dots > \gamma_{(N_t)}$, and 0 otherwise. We utilize L_t out of N_t of them, and choose the maximum

---

[1]The practical implications of this statement for antenna selection algorithms will be discussed in Sec. IV.

of all possible combinations. The desired $\gamma_{\rm bound}$ can be easily written in terms of the ordered SNRs as

$$\gamma_{\rm bound} = \sum_{i=1}^{L_t} \gamma_{(i)} \qquad (6)$$

3.2 Statistics of the SNR

The statistics of $\gamma_{\rm bound}$ can be derived from Eqs. 5 and 6. Mathematically, this problem is equivalent to computing the SNR for H-S/MRT with a single receive antenna, but with Nakagami channel statistics.[2] Consequently, the simple and elegant techniques for analyzing H-S/MRC with single-transmit-antenna in Rayleigh fading channels [15] cannot be used anymore. On the other hand, the available techniques for H-S/MRC in Nakagami-fading [2], [1] do not lend themselves easily to computer implementation. We are thus using a new approach, that also exploits the fact that in our case, the degrees of freedom (i.e. the number of antenna elements) can only take on integer values.

Since we are computing the sum of random variables, computing the characteristic function suggests itself naturally. We can write it as

$$\Phi(j\nu) = \frac{N_t!}{\Gamma(N_r)^{N_t}} \int_0^\infty d\gamma_{(1)} \gamma_{(1)}^{N_r-1} e^{-\gamma_{(1)}} e^{-j\nu\Xi(L_t-1)\gamma_{(1)}}$$

$$\int_0^{\gamma_{(1)}} d\gamma_{(2)} \gamma_{(2)}^{N_r-1} e^{-\gamma_{(2)}} e^{-j\nu\Xi(L_t-2)\gamma_{(2)}} \ldots$$

$$\int_0^{\gamma_{(N_t-1)}} d\gamma_{(N_t)} \gamma_{(N_t)}^{N_r-1} e^{-\gamma_{(N_t)}} e^{-j\nu\Xi(L_t-N_t)\gamma_{(N_t)}}$$

where $\Xi(x)$ is the Heaviside step function

$$\Xi(x) = \begin{cases} 1 & \text{if} \quad x \geq 0 \\ 0 & \text{otherwise} \end{cases} . \qquad (7)$$

In the following, we abbreviate the expression $j\nu\Xi(L_t-i)$ as a_i, dropping the dependence on ν for notational convenience. This multiple integral can be shown to result in a polynomial, whose coefficients can be computed analytically by a finite recursion with N_t iteration steps.

The crucial step of our proposed technique is now to recognize that an expression of the form

$$\left(d + \sum_p \exp(-b_p x) \wp(p, x) \right) \qquad (8)$$

where $\wp(p, x)$ is a polynomial in x whose coefficients may depend on p, retains its basic structure when integrated between 0 and y. Thus, the first $N_t - 1$ integrations can be written in an iterative fashion.

Specifically, let us write the integrand for the first integration (i.e. $q = 0$) as

$$\gamma_{(1)}^{N_r-1} \exp\left(-\gamma_{(1)}\right) I^{(0)} \qquad (9)$$

and quite generally denote the result of the qth integration as $I^{(q)}$, where superscript (q) indexes the number of performed integrations. The integral $I^{(q)}$ has the form

$$I^{(q)} = d^{(q)} + \sum_{p=1}^q e^{-b_p^{(q)}\gamma_{(N-q)}} \sum_{k=0}^{(q-p+1)(N_r-1)} c_{p,k}^{(q)} \gamma_{(N-q)}^k \qquad (10)$$

with initial condition

$$d^{(0)} = 1 \quad b_p^{(0)} = 0 \quad c_{p,k}^{(0)} = 0 \qquad (11)$$

It can be shown [9] that the central quantities $d^{(q)}, b_p^{(q)}$, and $c_{p,k}^{(q)}$ are given by recursion relations

$$b_p^{(q+1)} = b_p^{(q)} + a_{N_t-q} \qquad \text{for} \qquad 1 \leq p \leq q; \qquad (12)$$

$$\widehat{c}_{p,k}^{(q)} = c_{p,k-(N_r-1)}^{(q)} \qquad (13)$$

for $(q-p+2)(N_r-1) \geq k \geq (N_r-1)$ and 0 otherwise;

$$d^{(q+1)} = d^{(q)} \frac{(N_r-1)!}{(a_{N_r-q})^{N_r}} + \sum_{p=1}^q \sum_{t=0}^{(q-p+2)(N_r-1)} \frac{t!\widehat{c}_{p,t}^{(q)}}{(b_p^{(q+1)})^{t+1}} \qquad (14)$$

$$c_{p,k}^{(q+1)} = -\sum_{t=0}^{(q-p+2)(N_r-1)-k} \frac{\widehat{c}_{p,k+t}^{(q)}}{(b_p^{(q+1)})^{N_r}} \frac{(k+t)!}{k!} \qquad (15)$$

for $1 \leq p \leq q$ and

$$c_{p,k}^{(q+1)} = -\frac{d^{(q)}}{(b_p^{(q+1)})^{N_r-k}} \frac{(N_r-1)!}{k!} \qquad (16)$$

for $p = q + 1$.

The characteristic function of the $\gamma_{\rm bound}$ is finally given as

$$\Phi(j\nu) = \frac{N_t!}{\Gamma(N_r)^{N_t}} \Big[d^{(N_t)}(N_r-1)!a_1^{-N_r} + \qquad (17)$$

$$\sum_{p=1}^{N_t-1} \sum_{t=0}^{(N_r-p+1)(N_r-1)} \widehat{c}_{p,t}^{(N_r)} t! (b_p^{(N_t)})^{-(t+1)} \Big] .$$

Note that this is the characteristic function $\Phi(j\nu)$, where the coefficients $d^{(N_t)}$, $\widehat{c}_{p,t}^{(N_r)}$, and $b_p^{(N_t)}$ depend on $j\nu$.

In principle, an analytic inversion of the characteristic function would be possible, giving the probability density function of the SNR $p_{\gamma_{\rm bound}}(\cdot)$ in closed form. However, due to the existence of fast Fourier inversion techniques [11], numerical inversion is convenient and fast.

3.3 Symbol Error Probability and Capacity

Computation of the SEP can be done by the classical method of averaging the "instantaneous SEP" (i.e. SEP for one given channel realization) over the statistics of the SNR. Alternatively, we can compute the SEP via the characteristic function; for the explicit equations, see [12].

For a capacity point of view, the whole system between encoder and decoder can be viewed as an effective scalar flat fading channel that is characterized by γ_{bound}. The capacity for each channel realization is thus given by

$$C(\gamma_{bound}) = \log_2(1 + \bar{\Gamma}\gamma_{bound}). \qquad (18)$$

where $\bar{\Gamma}$ is the average SNR.

Using standard techniques for functions of one random variable [10], the pdf of the capacity becomes

$$p_C(C) = 2^C \frac{\ln(2)}{\bar{\Gamma}} p_{\gamma_{tot}}\left(\bar{\Gamma}(2^C - 1)\right). \qquad (19)$$

4 Results

In this section, we present results from our computations and discuss the influence of the number of available, and actually chosen, antennas on the system performance. If not stated otherwise, we will use the following system parameters: $\bar{\Gamma} = 20$ dB, $N_r = 2$. $N_t = 8$. For the bit error probability computations, we use minimum shift keying $\pi/4$–DQPSK since it is commonly used in mobile radio systems.

Figure 2 shows the cumulative distribution function (cdf) of the capacity for different numbers of selected antennas, L_t. The exact curve was computed by Monte Carlo (MC) simulations, the upper and lower bounds were computed by the analytical method described in Sec. 3. We note that upper and lower bound are 1 bit/s/Hz apart, except for the case $L_t = 1$, where they coincide and agree with the exact curve.

Apart from the bounds and the exact curves (computed by MC simulations), we also exhibit the cdf of the capacity when a suboptimum antenna selection criterion is used. This criterion works the following way: we transmit from a single antenna, $i = 1$, and determine the SNR that can be obtained at the receiver with MRC. Then, we transmit from the next antenna, $i = 2$, and determine again the SNR with MRC, and so on. Then, the L_t antennas that resulted in the best SNR are chosen. This can also be interpreted as optimizing γ_{bound} instead of λ_{max}. The advantage of this technique is that the determination of the "optimum" antennas is much simpler than if we have to make a full search among all possible antenna combinations. Furthermore, the loss in performance is less than 0.05 bits/s/Hz.

Figure 3 shows the increase of the mean capacity and 5% outage capacity as a function of the number of
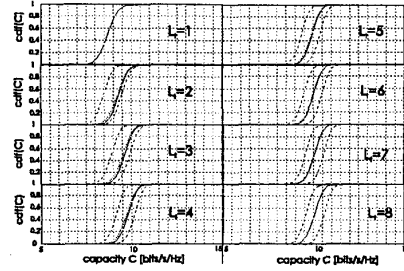

Figure 2. Cdf of the capacity: lower bound (left dashed curves), upper bound (right dashed curves), exact (solid curves), and exact with the use of the simplified selection criterion (solid). $N_t = 8$, $N_r = 8$, $SNR = 20dB$.
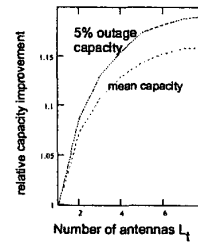

Figure 3. Capacity increase of the 5% outage capacity and the mean capacity compared to $L_t = 1$ when having several active antennas at the transmitter. $N_t = 8$, $N_r = 2$, $SNR = 20dB$.

selected antennas. We see that increasing that number from 1 to 2 gives about the same gain as increasing from 2 to 8. It seems thus reasonable to use only 2 or 3 selected antennas, giving large cost savings while reducing the performance only very little.

Figure 4 shows the downlink BER of $\pi/4$–DQPSK as a function of the mean SNR $\bar{\Gamma}$ for different number of selected antennas L_t. Again, we observe a big improvement going from $L_t = 1$ to $L_t = 3$ (3 dB @ $BER = 10^{-3}$), while the gain going from $L_t = 3$ to $L_t = 8$ is only an additional 1.5 dB.

Generally, the achieved capacities are much lower than those usually associated with MIMO systems. The difference is due to the restriction of the possible structure of the transmitter and receiver. Specifically, we allow only a scalar coder, and distinguish the signals at the different antennas only by linear weights, not by different codes at each antenna. Comparisons with truly optimum MIMO systems show that with appropriate (space-time) coding, an outage capacity of 16 bits/s/Hz is possible for $L_t = 8$, $N_r = 2$ [8]. The difference to the 10 bits/s/Hz obtained with the linear system is the price for backward compatibility and
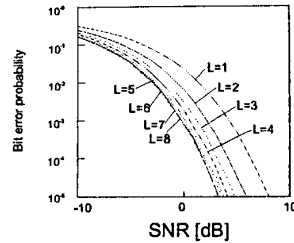
Figure 4. Bit error probability as a function of SNR for Gray-coded $\pi/4$-DQPSK as modulation format

greater simplicity. We also note that the increase in capacity slows down as we increase L_t, but shows no sharp discontinuity as L_t increases beyond $N_r = 2$. This is due to the fact that we use linear transmitters and receivers, so that every gain in SNR readily translates into a gain in capacity.

5   Summary and conclusions

We have investigated reduced-complexity wireless systems with transmit and receive diversity. The complexity reduction is achieved by using H-S/MRT on one link end, and MRC at the other. We note that for MRT and H-S/MRC, the results are equally applicable. Since the transceiver structure employs only weighted versions of the same signals, such a system is fully compatible with existing mobile radio systems, while the use of multiple antennas at both transmitter and receiver results in a high degree of diversity. The H-S/MRT(C) offers advantages when a large number of transmit antennas is available, but the number of RF chains should be limited. By choosing the best L_t out of N_t antennas, little signal quality is lost compared to the full-complexity version, while drastically reducing the involved hardware expenses. We have seen that for a practically useful example ($N_t = 8, N_r = 2$, $SNR = 20$ dB), $L_t = 2$ to 3 is a good compromise between hardware expense and performance.

References

[1] M. S. Alouini and M. K. Simon. Application of the Dirichlet transformation to the performance evaluation of generalized selection bomining over Nakagami-m fading channels. *J. Comm. Networks*, 1:5–13, 1999.

[2] M. S. Alouini and M. K. Simon. Performance of coherent receivers with hybrid SC/MRC over Nakagami-m fading channels. *IEEE Trans. Vehicular Techn.*, 48:1155–1164, 1999.

[3] J. B. Andersen. Antenna arrays in mobile communications: gain, diversity, and channel capacity. *IEEE Ant. Prop. Mag.*, 12–16, April 2000.

[4] G. J. Foschini and M. J. Gans. On limits of wireless communications in fading environments when using multiple antennas. *Wireless Personal Comm.*, 6:311–335, 1998.

[5] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge Univ. Press, 1985.

[6] N. Kong and L. B. Milstein. Combined average SNR of a generalized diversity selection combining scheme. In *Proc. ICC*, 1556–1560, 1998.

[7] T. K. Y. Lo. Maximum ratio transmission. In *Proc. ICC*, 1310–1314, 1999.

[8] A. F. Molisch, M. Z. Win, and J. H. Winters. Capacity of MIMO systems with antenna selection. In *Proc. ICC*, in press, 2001.

[9] A. F. Molisch, M. Z. Win, and J. H. Winters. Reduced-complexity wireless systems with transmit and receive diversity. *IEEE JSAC*, to be submitted, 2001.

[10] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. Tokyo, McGraw Hill, 1965.

[11] W. H. Press and et al. *Numerical Recipes*. Cambridge University Press, 1996.

[12] M. K. Simon and M.-S. Alouini. A unified approach to the performance analysis of digital communication over generalized fading channels. *Proc. IEEE*, 86(9):1858–1877, Sept. 1998.

[13] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths. *IEEE Trans. Comm.*, 47:199 –207, 1998.

[14] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Trans. IT*, 44:744–765, 1998.

[15] M. Z. Win and J. H. Winters. Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading. *IEEE Trans. Commun.*, 47(12):1773–1776, Dec. 1999.

[16] M. Z. Win and J. H. Winters. Analysis of hybrid selection/maximal-ratio combining of diversity branches with unequal SNR in Rayleigh fading. In *Proc. VTC*, 215–220, 1999.

[17] M. Z. Win and J. H. Winters. Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio combining in Rayleigh fading. *IEEE Trans. Commun.*, 2001. to appear.

[18] J. H. Winters. On the capacity of radio communications systems with diversity in Rayleigh fading environments. *IEEE J. Selected Areas Comm.* 5: 871–878, 1987.

[19] J. H. Winters. The diversity gain of transmit diversity in wireless systems with Rayleigh fading. In *Proc. ICC*, 1121–1125, 1994.

[20] A. Wittneben. A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation. In *Proc. ICC*, 1630–1634, 1993.