SIMPLE SPATIAL MULTIPLEXING BASED ON IMPERFECT CHANNEL ESTIMATES

Cristoff Martin, Svante Bergman, Björn Ottersten*

Dept. of Signals, Sensors & Systems
Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden

ABSTRACT

Techniques for communication over flat multi-input, multi-output (MIMO) channels are well established when either perfect channel state information or no channel state information is available at the transmitter. However, communication over channels where the transmitter has access to partial or imperfect information has received less attention. If exploited, such information could improve system performance. In this paper, we propose a simple system design scheme, that approximately maximizes the data rates of MIMO communication systems where imperfect channel estimates are available at the transmitter. The algorithm is computationally attractive and by taking the uncertainty of the channel estimates into account in the design gains compared with systems not exploiting this information can be demonstrated.

1. INTRODUCTION

Communication systems employing multiple transmit and receive antennas offer a dramatic increase in achievable data rates compared with single antenna alternatives [1, 2]. However, there are still some open issues in order to design a well optimized system in practice. For example, while well established techniques exist for cases where either perfect channel information or no channel information is available at the transmitter, see e.g. [2, 3], cases where partial or imperfect channel estimates are available are not so well understood.

Recently, some schemes have been developed that take advantage of partial channel information at the transmitter, see e.g. [4]. In these schemes, the transmitted data is optimized in order to minimize the error rate at the receiver given a certain bit rate. While this is desirable in some situations, in others it might be more attractive to, given a certain design bit error rate (BER), maximize the data rate of the system. When the channel is perfectly known, a practical solution to this optimization problem is spatial loading [2, 5]. Here, parallel, non-interfering spatial channels are first created using a linear transformation and then bit constellations and power are allocated for the different "spatial carriers" in order to maximize the data rate given some quality constraint on the received data, much in analogy with adaptive loading techniques commonly used in discrete multi-tone (DMT) systems [6].

In this preliminary work, we attempt take this idea one step further and generate a simple spatial loading scheme based on imperfect transmitter channel information. In contrast to the case where the channel is perfectly known at the transmitter and the case of DMT with imperfect channel estimates [7], it is not possible to create non-interfering parallel channels. This complicates the system design and increases the complexity of the receiver. Herein we present a simple and computationally efficient approximative method for the system design that is able to capture some of the potential gain in the available channel state information.

*The support of LIASIT (www.liasit.lu) and SES GLOBAL S.A., Luxembourg to this work is gratefully acknowledged. The author can be contacted at cristoff@s3.kth.se.

2. SYSTEM MODEL

The MIMO communication system considered consists of n_t transmitters communicating with a terminal with n_r receivers over a flat fading channel. A transmission over the channel is modeled through a channel matrix, $\boldsymbol{H} \in \mathbb{C}^{(n_\text{t} \times n_\text{r})}$, where the elements model the attenuations and phase shifts between the various transmitter, receiver pairs. The output, $\boldsymbol{x} \in \mathbb{C}^{n_\text{r}}$, resulting from a single usage of the channel to transmit the data $\boldsymbol{s} \in \mathbb{C}^{n_\text{t}}$, is then found as,

$$\boldsymbol{x} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{n} \in \mathbb{C}^{n_\text{r}}$ is additive white complex Gaussian noise. In order to make the required output power of the different transmission schemes presented in the paper comparable, the signal to noise ratio (SNR) is defined as the quotient between the power of the received data signal, $\operatorname{E}|\boldsymbol{H}\boldsymbol{s}|^2$, and the received noise power, $\operatorname{E}|\boldsymbol{n}|^2$, when equally powered uncorrelated data is transmitted, i.e when $\operatorname{E}\{\boldsymbol{s}\boldsymbol{s}^*\}$ is a scaled identity matrix. Without loss of generality, the received noise is normalized so that $\boldsymbol{n} \in \mathcal{CN}(0, \boldsymbol{I})$ and the channel matrix so that $\operatorname{E}\|\boldsymbol{H}\|_\text{F}^2 = n_\text{t} n_\text{r}$. The above definition and normalizations then give the SNR as $\operatorname{E}|\boldsymbol{s}|^2 = P$.

In what follows, the receiver is assumed to have perfect knowledge of $\boldsymbol{H}$ when it tries to detect the transmitted symbols. However, the channel knowledge at the transmitter is not necessarily perfect. There are several possibilities for the transmitter channel estimate to be in error. For example, the channel may have changed during the delay between the estimation and the usage of the estimates or the channel estimate may have been obtained via a low rate feedback channel not able to provide highly accurate estimates.

To model the uncertainty in the channel state information, we assume that the transmitter has access to an imperfect channel estimate $\hat{\boldsymbol{H}}$. This estimate relates to the true channel $\boldsymbol{H}$, which is available at the receiver, as,

$$\boldsymbol{H} = \hat{\boldsymbol{H}} + \tilde{\boldsymbol{H}}, \tag{2}$$

where $\tilde{\boldsymbol{H}}$ models the errors in the channel estimate. The elements of $\tilde{\boldsymbol{H}}$ are independently drawn from a complex Gaussian distribution. If this model fits reality well or if more elaborate modeling is required remains to be seen. Note from the normalizations above that the variance of the elements of $\tilde{\boldsymbol{H}}$, $\sigma_{\boldsymbol{H}}^2$, is normalized so that $0 < \sigma_{\boldsymbol{H}}^2 < 1$ where zero variance implies perfect channel knowledge, $\boldsymbol{H} = \hat{\boldsymbol{H}}$, and unit variance implies no knowledge, $\boldsymbol{H} = \tilde{\boldsymbol{H}}$, at the transmitter.

3. SPATIAL MULTIPLEXING AND LOADING

The basic idea behind bit and power loading algorithms is to optimize the transmitted vector data so that the available transmit power is used in the most efficient manner. For the system (1) the optimal transmit data design when the channel is known at the transmitter is, from a capacity [1] viewpoint, to diagonalize $\boldsymbol{H}$, creating parallel spatial channels, using a singular value decomposition and to transmit Gaussian distributed data symbols with power allocated using the well known water-filling solution. This way, transmit power is allocated to the directions where it is the most useful, i.e. the less attenuation in a direction the more power is allocated and the higher data rates can be supported.

For practical systems, using Gaussian distributed symbols is not an option. Instead the transmitted data symbols belong to some finite alphabet, resulting in an optimization problem where different constellation sizes and transmit powers are allocated to different directions in order to satisfy some constraint on the quality of the received data and to maximize data rates. Here, we term this type of technique spatial loading and in the subsections below such schemes are discussed and presented for different scenarios. In all cases we attempt to optimize the transmit data rate, given some design uncoded BER and transmit power. For a general spatial loading system it is natural to introduce some structure in order to characterize the spatial properties of the channel and to simplify the following adaptive loading. To that end, let $\boldsymbol{s} = \boldsymbol{V}_{\text{Tx}}\boldsymbol{P}^{1/2}\boldsymbol{d}$ and $\boldsymbol{y} = \boldsymbol{U}_{\text{Rx}}\boldsymbol{x}$, where $\boldsymbol{V}_{\text{Tx}}$ and $\boldsymbol{U}_{\text{Rx}}$ are unitary matrices characterizing the directivity of the spatial loading system, $\boldsymbol{P}$ is a diagonal matrix defining the power loading in the different directions, $\boldsymbol{y}$ is the received data to be considered by the detector and $\boldsymbol{d}$ is the transmitted symbols. Based on these definitions, the effective system between transmitter and receiver can be modeled as,

$$\boldsymbol{y} = \boldsymbol{U}_{\text{Rx}}(\boldsymbol{H}\boldsymbol{V}_{\text{Tx}}\boldsymbol{P}^{1/2}\boldsymbol{d} + \boldsymbol{n}) = \boldsymbol{H}'\boldsymbol{P}^{1/2}\boldsymbol{d} + \boldsymbol{n}', \qquad (3)$$

where $\boldsymbol{H}'$ is the effective channel matrix and $\boldsymbol{n}'$ is the effective noise of the MIMO channel between $\boldsymbol{d}$ and $\boldsymbol{y}$. Note that since $\boldsymbol{U}_{\text{Rx}}$ is unitary the effective noise is still white complex Gaussian, each element of variance one. Also, since $\boldsymbol{U}_{\text{Rx}}$ and $\boldsymbol{V}_{\text{Tx}}$ are invertible and the distributions of $\boldsymbol{n}$ and $\boldsymbol{n}'$ are identical the system (3) is equivalent to (1). The transmitted symbols are considered uncorrelated, with zero mean and normalized to variance one. The normalizations from Section 2 then implies that $\boldsymbol{P}$ should be chosen such that $\operatorname{Tr}\boldsymbol{P} = P$.

The focus of this preliminary study is on the uncertainty aspect of the channel, not on the receiver algorithm or bit loading scheme being used. In order to simplify the interpretation of the results and discussion below a maximum likelihood (ML) detector is employed in all cases and the well known Hughes-Hartogs algorithm [8, 6] has been selected for the bit and power loading. While these choices may be too computationally demanding for practical implementations, they simplify the derivation and presentation. For reference, in principle, the Hughes-Hartogs adaptive loading algorithm for parallel channels consists of the following steps,

1. Try to increase the constellation size by one for all the symbols of $\boldsymbol{d}$, one at the time. Add enough power so that the bit error constraint is not violated.
2. Increase the constellation size and allocate power to the symbol requiring the least additional power in the previous step.
3. Repeat until the available power is insufficient to add more bits given the error constraint. If desirable, any remaining power can be spread over the bits to improve the error rate performance of the wireless link.

3.1. Perfect Channel Knowledge

When the transmitter has perfect channel knowledge, a spatial loading scheme can be introduced based on a singular value decomposition of the channel matrix [2, 5]. Let $\boldsymbol{H} = \boldsymbol{U}_{\boldsymbol{H}}\boldsymbol{\Lambda}_{\boldsymbol{H}}\boldsymbol{V}_{\boldsymbol{H}}$ be the singular value decomposition of $\boldsymbol{H}$. The diagonal matrix $\boldsymbol{\Lambda}_{\boldsymbol{H}}$ contains the ordered singular values of $\boldsymbol{H}$, the unitary matrices $\boldsymbol{U}_{\boldsymbol{H}}$ and $\boldsymbol{V}_{\boldsymbol{H}}$ contains the corresponding left and right singular vectors.

The system is then designed as follows, the directive matrices at the transmitter and receiver are selected as $\boldsymbol{V}_{\text{Tx}} = \boldsymbol{V}_{\boldsymbol{H}}^*$ and $\boldsymbol{U}_{\text{Rx}} = \boldsymbol{U}_{\boldsymbol{H}}^*$. Note that the channel between $\boldsymbol{d}$ and $\boldsymbol{y}$ is diagonalized and an equivalent system $\boldsymbol{y} = \boldsymbol{\Lambda}_{\boldsymbol{H}}\boldsymbol{P}^{1/2}\boldsymbol{d} + \boldsymbol{n}'$ is formed, i.e. there are $\min(n_{\text{r}}, n_{\text{t}})$ non-interfering channels between the transmitted and the received data.

Since the above step creates an equivalent system of $\min(n_{\text{r}}, n_{\text{t}})$ parallel channels, each characterized by an SNR given by the singular values of $\boldsymbol{H}$, bit and power loading using the algorithm above is straightforward.

3.2. No Channel Knowledge

For the case when the transmitter has no knowledge of the channel it is not possible to optimize the directivity of the transmitting array. Here, we assume that a non-line of sight system is considered where the antenna elements are sufficiently separated so that the elements of $\boldsymbol{H}$ can be considered independent Rayleigh fading.

For this type of system several techniques such as BLAST [9] or more sophisticated space-time coding schemes [3] have been designed. While elaborate techniques are required for efficient detection of the transmitted symbols and coding is required for optimal performance we here only consider a system where the transmitter transmits uncoded symbols, $\boldsymbol{y}$, with equally distributed power, $\boldsymbol{P} = P/n_{\text{t}}\boldsymbol{I}$, and the receiver uses an ML-detector to estimate the transmitted symbols.

3.3. Imperfect Channel Knowledge

When perfect channel knowledge is not available at the transmitter the problem becomes more complicated. For example, it is no longer possible to completely diagonalize the channel and the parallel data streams are interfering at the receiver. Here, we propose a simple sub-optimal scheme for optimizing the data rates and exploit available transmitter channel knowledge.

Let the singular value decomposition of $\hat{\boldsymbol{H}}$ be written $\hat{\boldsymbol{H}} = \boldsymbol{U}_{\hat{\boldsymbol{H}}}\boldsymbol{\Lambda}_{\hat{\boldsymbol{H}}}\boldsymbol{V}_{\hat{\boldsymbol{H}}}$. The transmitter and receiver directive matrices, $\boldsymbol{V}_{\text{Tx}}$ and $\boldsymbol{U}_{\text{Rx}}$ are selected as $\boldsymbol{V}_{\text{Tx}} = \boldsymbol{V}_{\hat{\boldsymbol{H}}}^*$ and $\boldsymbol{U}_{\text{Rx}} = \boldsymbol{U}_{\hat{\boldsymbol{H}}}^*$ resulting in an efficient system,

$$\boldsymbol{y} = (\boldsymbol{\Lambda}_{\hat{\boldsymbol{H}}} + \tilde{\boldsymbol{H}}')\boldsymbol{P}^{1/2}\boldsymbol{d} + \boldsymbol{n}'. \qquad (4)$$

where $\tilde{\boldsymbol{H}}' = \boldsymbol{U}_{\text{Rx}}\tilde{\boldsymbol{H}}\boldsymbol{V}_{\text{Tx}}$ has the same distribution as $\tilde{\boldsymbol{H}}$ since $\boldsymbol{V}_{\text{Tx}}$ and $\boldsymbol{U}_{\text{Rx}}$ are unitary. While we do not claim that this choice of $\boldsymbol{V}_{\text{Tx}}$ and $\boldsymbol{U}_{\text{Rx}}$ is optimal it can be motivated in a number of ways. Firstly, given the channel knowledge at the transmitter, the received signal power for the first symbol in $\boldsymbol{d}$, $\operatorname{E}|\boldsymbol{U}_{\text{Rx}}\boldsymbol{H}\boldsymbol{V}_{\text{Tx}}\boldsymbol{P}^{1/2}\boldsymbol{d}_1|^2$ is maximized. Here, $\boldsymbol{d}_k$ is a vector where all entries are zero except for the kth entry which is identical to kth element of $\boldsymbol{d}$. Similarly the second element is transmitted in the orthogonal direction, in $\mathbb{C}^{n_{\text{t}}}$, to the first, that maximizes the received signal power and so on. Hence given this choice of directive matrices it is possible to ensure that data is transmitted in the directions where the receive conditions are likely to be the most favorable. Secondly, since $\boldsymbol{\Lambda}_{\hat{\boldsymbol{H}}}$ is diagonal and $\tilde{\boldsymbol{H}}'$ spatially white, the received power from the first transmitted symbol of $\boldsymbol{d}$ is maximized at the first received symbol of $\boldsymbol{y}$ while the received power at the other elements of $\boldsymbol{y}$ is minimized given the available channel information. Similarly, the received power resulting from the second symbol of $\boldsymbol{d}$ is concentrated to the second element of $\boldsymbol{y}$, and so on. Thus, while interference between the different spatial carriers is unavoidable, this choice of directive matrices in some sense minimizes it. Finally, for the case $\sigma_{\boldsymbol{H}}^2 \rightarrow 0$, this solution coincides with that of Section 3.1, and, if the elements of the true channel can be considered independent Rayleigh fading, the choice is also valid when $\sigma_{\boldsymbol{H}}^2 \rightarrow 1$, see Section 3.2.

In order to provide a practical spatial bit and power loading scheme it is necessary to be able compute the resulting BER for various constellation sizes and output powers efficiently. Since the communication channels are interfering this is complicated and for a computationally attractive scheme some approximations are necessary. Firstly, we assume that the design BER is chosen so low that more than one symbol error per received vector $\boldsymbol{y}$ is rare. Thus, when computing the error rates of the ML-detector for each of the transmitted symbols in $\boldsymbol{d}$ we assume that the other symbols have been correctly detected and subtracted from the received data, i.e error propagation is ignored. Note that this approximation can be expected to work better when there are more receive than transmit antennas, and in practice it is good to limit the number of spatial channels to $\min(n_{\mathrm{r}}, n_{\mathrm{t}})$ or less. Furthermore, it is assumed that the transmitted symbols have been Gray-encoded so that, given the low design bit error rate, the number of bit errors can be approximated by the number of symbol errors. These approximations, result in a simplified scenario where it suffices to compute the error probability of a single symbol, transmitted over a vector channel consisting of a combination of independent Rayleigh and Ricean fading elements. For this type of channels efficient techniques for computing the error probability for many types of constellations exist [10]. Note that since the variances of the vector channel elements are determined by $\sigma_{\boldsymbol{H}}^2$ the bit and power loading in the different directions are adapted to the uncertainty of the channel.

Using the selected $\boldsymbol{V}_{\mathrm{Tx}}$ and $\boldsymbol{U}_{\mathrm{Rx}}$ and the approximations above, performing the spatial loading can be performed in a simple and computationally efficient manner. While this algorithm is clearly suboptimal, note that for the cases when the channel information is almost perfect, or when the channel is almost completely unknown, our solution converges to well known results, see Sections 3.1 and 3.2.

4. NUMERICAL RESULTS AND ANALYSIS

Since an analytical analysis on the performance of the method proposed in the previous section would be difficult, a numerical analysis is performed. The simulations performed in this preliminary study were simplified by limiting the bit-loading algorithm to square M-QAM constellations, where $M = 2^{2b}$ and is b integer valued. Note that results such as those in [10] allow a larger selection of constellations which might be desirable.

To make the simulations efficient, the ML-detector at the receiver was implemented in the form of a sphere decoder [11]. For reasonably sized arrays the sphere decoding algorithm is on average very efficient, significantly shortening the time required to find the ML-solution compared with full search. Note that in the case of spatial loading the sphere decoding algorithm needs to be adapted to consider the different constellations received in the different elements of $\boldsymbol{d}$.

4.1. Ex 1: LOS channel with local scattering

Consider a line of sight scenario with significant scattering around the arrays. Provided that the antenna elements of the arrays are spaced too closely to be separated based on the line of sight component of the received data, but so well separated that the fading caused by the multi-path scatterers can be considered independent, a reasonable channel model is

$$\boldsymbol{H} = \boldsymbol{H}_{\mathrm{LOS}} + \boldsymbol{H}_{NLOS}. \tag{5}$$

Here, $\boldsymbol{H}_{\mathrm{LOS}}$ is a rank one matrix modeling the line of sight component, $\boldsymbol{H}_{\mathrm{NLOS}}$ is a random matrix where the elements are independently drawn from a zero-mean complex Gaussian source. In some scenarios the scattering components of the channel may be rapidly changing while the line of sight component and the statistical properties of the scattering component is changing more slowly. Hence, while keeping exact channel estimates at the transmitter may be difficult, maintaining estimates of $\boldsymbol{H}_{\mathrm{LOS}}$ and the variance of the elements of $\boldsymbol{H}_{\mathrm{NLOS}}$ could be feasible. This scenario fits well within the framework of our proposed method, $\hat{\boldsymbol{H}}$ then corresponds to $\boldsymbol{H}_{\mathrm{LOS}}$ and $\sigma_{\boldsymbol{H}}^2$ to the variance of the elements of $\boldsymbol{H}_{\mathrm{NLOS}}$.

Fig. 1 shows the bit rate as a function of the resulting bit error rate of a three by three antenna system where the line of sight component provides 80% of the received signal power, i.e. $\sigma_{\boldsymbol{H}}^2 = 0.2$. Two methods of system design are considered, one using only the line of sight rank one channel estimate in the design of the bit and power loading, the other using our proposed method. In both cases, any remaining transmit power after the bit loading is spread evenly over the spatial carriers in order to ensure SNRs of 15, 20 and 25 dB. By taking the unknown multi-path components into account the simple method proposed in Section 3.3 is capable of increasing the data rate, with significant gains at high SNRs.

To further illustrate the behavior of the algorithm, Table 1 shows how the transmitted bits are allocated as a function of $\sigma_{\boldsymbol{H}}^2$. As the resulting error rates vary greatly between the different channels this result gives no indication of resulting bit rates but is intended to illustrate how the bits are allocated to the different carriers. Note that when $\sigma_{\boldsymbol{H}}^2 \rightarrow 0$ signal power and loaded bits are concentrated in the direction given by the line of sight channel, while when $\sigma_{\boldsymbol{H}}^2 \rightarrow 1$ power and data are transmitted as evenly as possible in space given the restriction in constellation sizes. This intuitive result illustrates how this method, while simple, is capable of adapting the transmitted data according to the transmitter channel knowledge available.

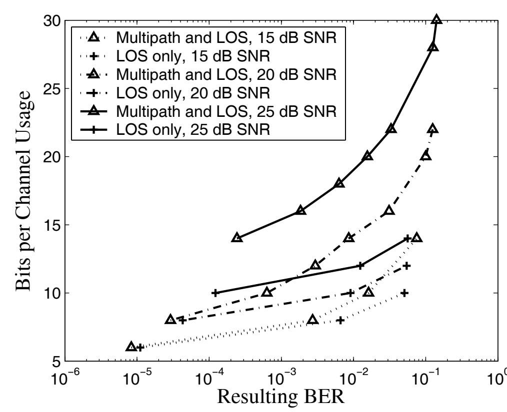

Fig. 1. Taking multi-path components into account vs. ignoring it, $\hat{\boldsymbol{H}}$ rank one, $\sigma_{\boldsymbol{H}}^2 = 0.2$, $n_{\mathrm{r}} = n_{\mathrm{t}} = 3$, SNR 15, 20, 25 dB.

4.2. Ex 2: Rayleigh Fading with Imperfect Channel Estimates

The elements of both $\hat{\boldsymbol{H}}$ and $\tilde{\boldsymbol{H}}$ are drawn independently from Gaussian distributions and consequently the channel $\boldsymbol{H}$ is Rayleigh fading. This could for example model an indoor non line of sight channel with well separated antenna elements at both arrays. In this example a 4 transmit, 6 receive antenna system designed for an SNR of 20 dB is considered.

σ_H^2	0.0	0.1	0.2	0.3	0.4	0.5	0.6	...	0.9	1.0
Chan. 1:	8	8	8	8	6	6	6	...	6	4
Chan. 2:	0	2	2	2	4	4	4	...	4	4
Chan. 3:	0	0	2	2	4	4	4	...	4	4
Chan. 4:	0	0	2	2	2	2	4	...	4	4

Table 1. Bit allocation, design SNR 20 dB, design BER 0.003.

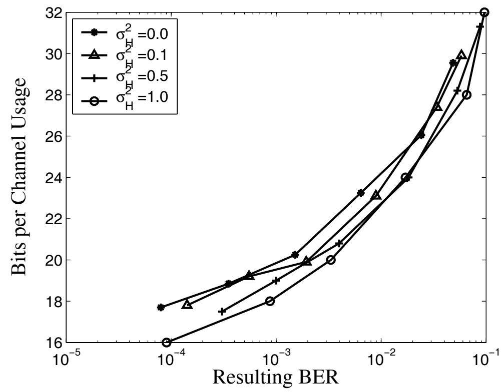

Fig. 2. Independent Rayleigh fading, $n_r = 6$, $n_t = 4$, SNR 20 dB.

Fig. 2 shows bit rate versus resulting BER at various degree of uncertainty of the channel estimate, σ_H^2. Any remaining power after the bit loading is spread evenly over the carriers in order to make the comparison fair. This figure clearly illustrates the capability of the proposed method to connect well known results for the design of system with perfect channel knowledge at the transmitter, $\sigma_H^2 = 0$ to designs used when no channel knowledge, $\sigma_H^2 = 1$, is available.

Finally, the performance of the approximations leading up to the bit loading algorithm is evaluated. Fig. 3 shows the resulting BER as a function of the design BER. In order to be able to evaluate the performance of the approximations, the remaining power after the bit allocation is not allocated to any spatial carrier. As the approximations used in deriving the proposed method are underestimating the probability of error, the design BER is lower than the resulting BER. While this means that the design BER must be chosen lower than that required by the system, the proposed method achieves balancing in the loading of the different spatial channels. Furthermore, notice that as σ_H^2 gets close to zero, the approximations improve drastically. This is expected as small σ_H^2 means less interference and error propagation between the channels and the approximation to ignore these effects is improved.

5. CONCLUSIONS

Herein, a simple and computationally efficient method for data rate optimization of practical MIMO communication systems with imperfect channel information at the transmitter was introduced. Results from simulations indicate that the algorithm provides gains compared with methods that either do not take available information into account, or completely trust the information that is only partly correct. Also, for cases where the transmitter has perfect or no channel knowledge, the scheme converges to well known solutions, connecting spatial multiplexing techniques with transmission schemes used over unknown independent Rayleigh fading channels.

ACKNOWLEDGMENTS

We would like to thank our colleague Joakim Jaldén for providing us with the sphere decoder used in our experiments.

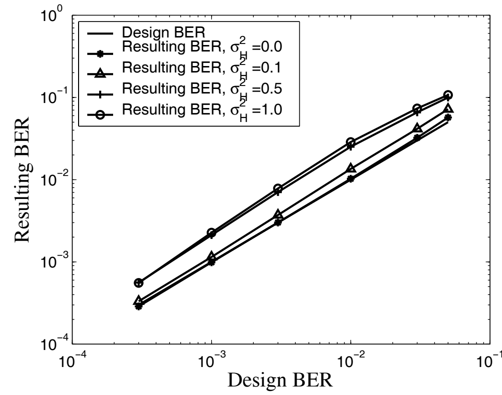

Fig. 3. Independent Rayleigh fading, error rate performance of the approximations, $n_r = 6$, $n_t = 4$, design SNR 20 dB.

6. REFERENCES

[1] Emre Telatar, "Capacity of multi-antenna Gaussian channels," *Technical Memorandum, Bell Laboratories (Published in European Transactions on Telecommunications, Vol. 10, No.6, pp. 585-595, Nov/Dec 1999)*, 1995.

[2] Gregory G. Raleigh and John M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 357–366, March 1998.

[3] Vahid Tarokh, Nambi Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, March 1998.

[4] George Jöngren, Mikael Skoglund, and Björn Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 48, pp. 611–627, March 2002.

[5] Xi Zhang and Björn Ottersten, "Power allocation and bit loading for spatial multiplexing in MIMO systems," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2003.

[6] John A.C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, pp. 5–14, May 1990.

[7] Achankang Leke and John M. Cioffi, "Multicarrier systems with imperfect channel knowledge," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, September 1998, vol. 2, pp. 549–553.

[8] Dirk Hughes-Hartogs, "Ensemble modem structure for imperfect transmission media," United States Patents, 4,679,227 (July 1987), 4,731,816 (March 1998), 4,833,706 (May 1998).

[9] Gerard J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, Autumn 1996.

[10] Mohamed-Slim Alouini and Andrea J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE J. Select. Areas Commun.*, vol. 47, no. 9, pp. 1324–1334, September 1999.

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