# Design of Multidimensional Mappings for Iterative MIMO Detection with Minimized Bit Error Floor

Nabil Sven Muhammad, Joachim Speidel

Institute of Telecommunications, University of Stuttgart, Pfaffenwaldring 47, D-70569 Stuttgart, Germany, {muhammad,speidel}@inue.uni-stuttgart.de

## Abstract

We present multidimensional mapping (MdM) as a generalized way how to assign input bits to symbol vectors for spatial multiplexing multiple-input, multiple-output (MIMO) transmission with iterative detection. Design goal is to minimize bit error floor by means of EXIT chart and pairwise error probability under the assumption of perfect a priori knowledge at the demapper. We propose an effective description of MdM based on block code generator matrices. The presented MdMs achieve minimal error floor and turbo cliff positions close to capacity limit.

## 1 Introduction

Utilization of multiple antennas at both transmitter and receiver allows high data rates as well as reliable communication. Multiple-input, multiple-output (MIMO) systems are therefore still a field of intensive research. In [1] the potential for enormous capacity gains was first adressed. High data rates can be achieved by spatial multiplexing, e.g. with the V-BLAST architecture [2].

As in most disciplines, a tradeoff between complexity and performance of a system has to be made. Approaching capacity inherently demands more sophisticated transmitter and to an even greater extend receiver schemes. One way to improve performance is iterative MIMO detection, often referred to as bit-interleaved coded modulation with iterative detection (BICM-ID) for MIMO. Unless the outer code is of low rate these schemes typically suffer from a serious error floor. Decreasing the error floor can be achieved by code doping, where a rate 1 inner recursive encoder is applied [3], [4]. As a drawback, an additional maximum a posteriori (MAP) decoder is required at the receiver. Alternatively, mappings optimized for BICM-ID can be used [5]. As opposed to Gray mapping, symbols differing in one bit should have a large Euclidean distance. This is justified by the availability of a priori information. No additional overhead is required and error floor is slightly decreased. Another way to decrease bit error rate (BER) is to employ an optimal (vector-wise) MAP demapper. For most MIMO scenarios this would be prohibitively complex, but close-to-optimum sphere decoding algorithms exist with lower complexity, e.g. [6], [7].

In this paper we examine the conjunction of optimized mappings with iterative MAP demapping for improved MIMO transmission. Multidimensional mapping (MdM) was proposed in [3] and brute-force search optimization was performed. We derive an effective description of MdMs based on block code generator matrices. Using EXIT chart [8] and bounds on pairwise error probability (PEP) for BICM [9], we propose optimized MdMs for several configurations, which allow for transmission with minimal error floor and turbo cliff positions close to capacity limit.

In Section 2 we describe the BICM spatial multiplexing transmitter, the channel model and the iterative receiver. Section 3 outlines MdMs and encoding using generator matrices. Design critria are derived in Section 4 and optimized MdMs are presented, which are verified by simulations results in Section 5, followed by the conclusion in Section 6.

## 2 System Model

### 2.1 Transmitter

The transmitter in Fig. 1 consists of a bit source with output sequence $\mathbf{u}$. These information bits are first convolutionally encoded by an outer encoder of rate $R_c$, then randomly bit-interleaved with an interleaving depth $S$. Input to the MdM is row vector $\mathbf{x} = (x_1, \ldots, x_{QM})$ comprising $QM$ bits. $M$ is the number of transmit (tx) antennas and $Q$ is the number of bits per complex symbol. The symbol column vector $\mathbf{s} = (s_1, \ldots, s_M)^T$ is transmitted, where each complex symbol $s_j, j \in \{1, \ldots, M\}$, has mean power $E_s$. $(.)^T$ denotes transposition. An MdM is defined by the relation $\mathbf{s} = \mu(\mathbf{x})$, with $\mu(.)$ being the multidimensional mapping function, and is examined more closely in Section 3.

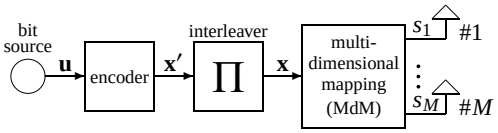


Fig. 1: MIMO transmitter with multidimensional mapper

## 2.2 Channel model

We use the widely accepted frequency flat fading MIMO equivalent baseband model, described by

$$\mathbf{r} = H\mathbf{s} + \mathbf{n} \quad . \tag{1}$$

$\mathbf{r} = (r_1, \ldots, r_N)^{\mathrm{T}}$ is the received symbol column vector, where $r_i$ is the component for antenna $i, i \in \{1, \ldots, N\}$, and $N$ is the number of receive (rx) antennas. The channel matrix is given by $H = (\mathbf{h}_1 \ldots \mathbf{h}_M)$, with $\mathbf{h}_j = \left(h_{1,j}, \ldots, h_{N,j}\right)^{\mathrm{T}}$. The impulse response $h_{i,j}$ from transmitter $j$ to receiver $i$ is modelled as a zero-mean, complex Gaussian random variable satisfying $\mathsf{E}\left\{\left|h_{i,j}\right|^2\right\} = 1$ (i.e., the channel is passive). All entries of $H$ are i.i.d., change with each channel use and are assumed to be perfectly known at the receiver. $\mathbf{n} = (n_1, \ldots, n_N)^{\mathrm{T}}$ is an additive noise column vector with components $n_i$ at receive antenna $i$, which are complex AWGN, each with zero-mean and variance $\sigma_{\mathrm{n}}^2$. This means that real and imaginary part of $n_i$ are Gaussian with variance $\sigma_{\mathrm{n}}^2/2 = N_0/2$ each. We consider $\mathbf{n}$ to be uncorrelated, i.e., $\mathsf{E}\left\{\mathbf{n}\mathbf{n}^{\mathrm{H}}\right\} = \sigma_{\mathrm{n}}^2 I_N$. $I_N$ is the identity matrix of size $N$ and $(.)^{\mathrm{H}}$ denotes conjugate transposition. We define the signal-to-noise ratio (SNR) as

$$\mathrm{SNR} = \frac{E_{\mathrm{b}}}{N_0} = \frac{M \cdot E_{\mathrm{s}} \cdot N}{R_{\mathrm{c}} \cdot M \cdot Q \cdot N_0} = \frac{E_{\mathrm{s}} \cdot N}{R_{\mathrm{c}} \cdot Q \cdot \sigma_{\mathrm{n}}^2}, \tag{2}$$

where $E_{\mathrm{b}}$ is the transmitted energy per information bit at the receiver and $N_0$ the noise power spectral density. Note that (2) is the SNR at the total receiver and not at one receiver antenna element.

## 2.3 Receiver

The iterative receiver is depicted in Fig. 2.

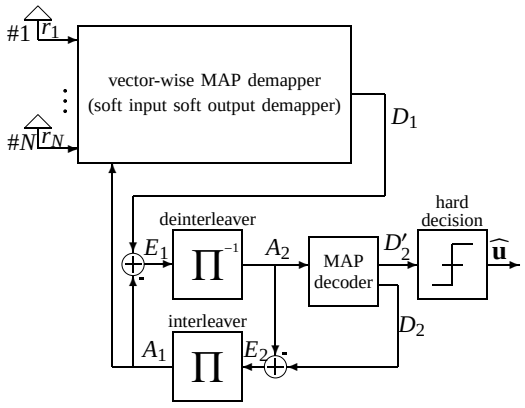


Fig. 2: Receiver structure for iterative detection

The vector-wise MAP demapper computes for each bit $x_k, k \in \{1, \ldots, QM\}$, a posteriori $L$-values $L_{D,1}(x_k|\mathbf{r}) \div D_1$ out of channel observation $\mathbf{r}$ and a priori $L$-values $\mathbf{L_{A,1}} = (L_{A,1}(x_1), \ldots, L_{A,1}(x_{QM}))^{\mathrm{T}}$, fed back from the outer MAP decoder. Subtraction of $L_{A,1}(x_k) \div A_1$ yields extrinsic $L$-value $L_{E,1}(x_k, \mathbf{r}) \div E_1$ as [10]:

$$L_{E,1}(x_k, \mathbf{r}) = \ln \frac{\sum\limits_{\hat{\mathbf{x}} \in \mathbb{X}_{k,1}} p(\mathbf{r}|\mu(\hat{\mathbf{x}})) \cdot \exp\left(\hat{\mathbf{x}}_{[k]} \cdot \mathbf{L}_{\mathbf{A},\mathbf{1}[k]}\right)}{\sum\limits_{\hat{\mathbf{x}} \in \mathbb{X}_{k,0}} p(\mathbf{r}|\mu(\hat{\mathbf{x}})) \cdot \exp\left(\hat{\mathbf{x}}_{[k]} \cdot \mathbf{L}_{\mathbf{A},\mathbf{1}[k]}\right)}, \tag{3}$$

where $\mathbb{X}_{k,b}$ is the set of all row vectors $\hat{\mathbf{x}}$ for which $\hat{x}_k = b, b \in \{0, 1\}$, and $\hat{\mathbf{x}}_{[k]} \cdot \mathbf{L}_{\mathbf{A},\mathbf{1}[k]}$ is the scalar product of both vectors without the $k$th element. $A_1$, $D_1$ and $E_1$ are shorthand notations. Other $L$-values occuring in Fig. 2 are defined similarly. The probability density functions $p(\mathbf{r}|\mu(\hat{\mathbf{x}}))$ are multidimensional Gauss-functions:

$$p(\mathbf{r}|\mu(\hat{\mathbf{x}})) = \frac{1}{(\pi N_0)^M} \exp\left(-\frac{\|\mathbf{r} - H\mu(\hat{\mathbf{x}})\|^2}{N_0}\right). \tag{4}$$

After deinterleaving, $E_1$ is used as a priori information $A_2$ for the MAP decoder, which outputs a posteriori $L$-values $D_2$ of coded bits. In the first pass through the demapper the a priori $L$-values $A_1$ are set to zero, while in the following iterations $E_2 = D_2 - A_2$ are interleaved and the resulting a priori information $A_1$ is fed back. After a certain number of iterations, a hard decision on the a posteriori $L$-values $D_2'$ of the information bits yields the estimates $\widehat{\mathbf{u}}$ on the transmitted data.

To reduce complexity of computation of (3), which is exponential with $QM$, it would be reasonable to apply suboptimal sphere decoding for practical implementations [7]. In this paper we use optimal MAP demapping to obtain lower bounds on BER.

# 3 Multidimensional Mappings

## 3.1 Principle

We first consider conventional symbol-wise mapping, which is applied, e.g., in the V-BLAST architecture [2]. We write the input vector as $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_M)$, with $\mathbf{x}_j = (x_{j1}, \ldots, x_{jQ}), j \in \{1, \ldots, M\}$. For a given signal constellation, the complex symbol $s_j$, transmitted over antenna $j$, is determined by $Q$ bits in vector $\mathbf{x}_j$ only. Real and imaginary part of $s_j$ are obtained by a two-dimensional (2d) mapping, denoted as $s_j = \mu_{2\mathrm{d}}(\mathbf{x}_j)$. We can thus write the transmit vector as $\mathbf{s} = \mu(\mathbf{x}) = (\mu_{2\mathrm{d}}(\mathbf{x}_1), \ldots, \mu_{2\mathrm{d}}(\mathbf{x}_M))^{\mathrm{T}}$. A linear detector applies zero forcing (ZF) or minimum mean squared error (MMSE) equalization, followed by symbol-wise MAP demapping. Vector-wise MAP demapping is also applicable and is known to be the optimal detection scheme.

For the following example we define $\mu_{2\mathrm{d}}(.)$ to be a binary phase shift keying (BPSK) mapping function ($Q = 1$), i.e., $\mu_{2\mathrm{d}}(0) = -\sqrt{E_{\mathrm{s}}}$ and $\mu_{2\mathrm{d}}(1) = +\sqrt{E_{\mathrm{s}}}$. Although BPSK is a one-dimensional mapping, we maintain the notation $\mu_{2\mathrm{d}}(.)$. An MdM generalizes a mapping by allowing for any one-to-one correspondence from bit vector $\mathbf{x}$ to symbol vector $\mathbf{s} = \mu(\mathbf{x})$. However, we restrict ourselves to the case, where all symbols $s_j$ are drawn from the same modulation alphabet. Table I shows an MdM for $M = 2$ transmit antennas and BPSK symbols.

TABLE I: Example for MdM: $M = 2$ transmit antennas and BPSK, $Q = 1$

| $\mathbf{x}_1$ | $\mathbf{x}_2$ | $s_1$ | $s_2$ | $\mathbf{c}_1$ | $\mathbf{c}_2$ |
|---|---|---|---|---|---|
| 0 | 0 | $-\sqrt{E_s}$ | $-\sqrt{E_s}$ | 0 | 0 |
| 1 | 0 | $+\sqrt{E_s}$ | $+\sqrt{E_s}$ | 1 | 1 |
| 0 | 1 | $+\sqrt{E_s}$ | $-\sqrt{E_s}$ | 1 | 0 |
| 1 | 1 | $-\sqrt{E_s}$ | $+\sqrt{E_s}$ | 0 | 1 |

The multidimensional mapper uses the look-up table to allocate $QM$ input bits to a transmit symbol vector $\mathbf{s}$, e.g., $\mu((1,0)) = (+\sqrt{E_s}, +\sqrt{E_s})^T$. Thus, linear ZF or MMSE detection is not possible anymore. The look-up table consists of $2^{QM}$ rows. For small $Q$ and $M$, the table is manageable. However, for larger $Q$ and $M$, its size gets very large, e.g., for $Q = 2$ and $M = 4$, we would have 256 rows. We thus derive an alternative description of MdM in the following subsection. Note that optimization of MdMs by brute-force search is literally intractable, since there exist $(2^{QM})!$ different possibilities to define the look-up table. With $Q = 2$ and $M = 4$, there are about $10^{505}$ different look-up tables.

### 3.2 MdM generation

In many cases, an MdM can be described by

$$\mathbf{s} = \mu(\mathbf{x}) = (\mu_{2d}(\mathbf{c}_1), \ldots, \mu_{2d}(\mathbf{c}_M))^T, \quad (5)$$

$$\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_M) = \mathbf{x} \cdot G \quad (6)$$

$$\mathbf{c}_j = (c_{j1}, \ldots, c_{jQ}), \quad j \in \{1, \ldots, M\}. \quad (7)$$

The generator matrix $G$ is square, non-singular and of dimension $QM$. Its entries are $\in \{0,1\}$. Additions and multiplications in (6) have to be performed in Galois Field GF(2). With (5) and (6) an MdM can be constructed by applying conventional symbol-wise mapping $\mu_{2d}(.)$ to the coded bits $\mathbf{c}$. For the example from Table I we conclude

$$G = G_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}. \quad (8)$$

For higher dimensions ($QM > 2$), it is not always possible to describe any MdM by (6). However, all optimized MdMs presented in Section 4 allow this description. In Table I, the coded bits $\mathbf{c}$ are also listed.

$\mu_{2d}(.)$ can now be regarded as a two-dimensional reference mapping. If $Q = 1$, we use the previous BPSK definition, if $Q = 2$, we define $\mu_{2d}(.)$ as the quadrature phase shift keying (QPSK) Gray mapping and if $Q = 4$, we will make use of two different reference mappings shown in Fig. 3. One is 16-QAM (quadrature amplitude modulation) Gray mapping $\mu_{2d}(.)$. If we swap four symbols as indicated by the arrows in Fig. 3, we obtain a modified mapping $\mu'_{2d}(.)$, which is per definition not a Gray mapping anymore. $\alpha = 2/5 \cdot E_s$ to obtain mean symbol power $E_s$. Note that with our notation the leftmost bit corresponds to the least significant bit.

Fig. 4 depicts an alternative scheme to Fig. 1 for MdM using eqs. (5)-(7). Multiplication of the input bits $\mathbf{x}$ with $G$ corresponds to linear block encoding

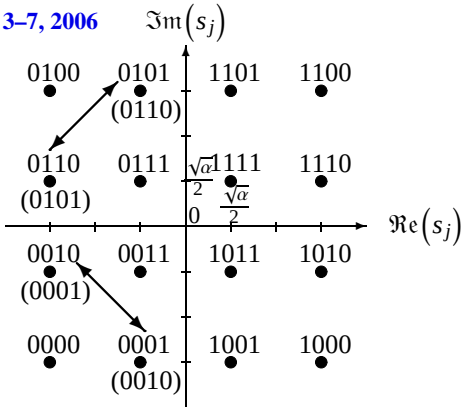


Fig. 3: Reference mappings for 16-QAM: Gray $\mu_{2d}(.)$ and modified mapping $\mu'_{2d}(.)$ (with changes in brackets)

with code rate 1. The encoded bits $\mathbf{c}$ are input to a conventional V-BLAST transmitter, where $Q$ bits are assigned to complex, i.e., two-dimensional symbols, which are then serial-to-parallel converted before transmission over the antennas.

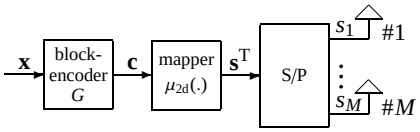


Fig. 4: Alternative scheme for MdM

At this point it is worth mentioning that also mappings for a single-antenna transmitter ($M = 1$) can be described by a reference mapping $\mu_{2d}(.)$ and an MdM table. With properly definined $\mu_{2d}(.)$, the generator matrix notation can be used as well. As an example, it can easily be verified that QPSK anti-Gray mapping corresponds to block encoding with $G_2$ in (8), followed by Gray mapping.

Finally, we introduce a short-hand notation that will simplify the description of the generator matrices $G$. Let $\mathbf{1}$ be a column vector of dimension $QM$, with all entries 1, and let $\overline{\mathbf{e}_i}$ be the $i$-th inverted unit column vector of dimension $QM$, i.e., a vector, with all entries 1, except the $i$-th position, where there is a 0. Then, $G_2$ in (8) can be written as $G_2 = (\mathbf{1}\ \overline{\mathbf{e}_2})$.

## 4 Design Criteria and MdM Optimization

Optimization of MdMs can be considered for two applications. First, if the receiver does not perform iterative decoding, i.e., the demapper does not use a priori knowledge, it is well known that the optimum MdM is symbol-wise Gray mapping. The generator matrix $G$ then is the identity matrix and the mapping functions $\mu_{2d}(.)$ as defined in the previous section are used. Note that an identity matrix is a non-singular matrix with maximum number of 0 entries.

Secondly, a receiver, which iterates over demapper and decoder, can utilize a priori information at the demapper. We focus on the case of perfect a priori information. This means that the outer decoder provides perfect knowledge about all bits $\mathbf{x}$ except the bit $x_k$ under consideration in (3). Utilizing this knowledge most efficiently corresponds to maximizing mutual information $I_{E1}(I_{A1}=1)$ in the EXIT chart. The higher $I_{E1}(I_{A1}=1)$, the closer the intersection of the transfer characteristics belonging to demapper and decoder comes to the desired upper right corner of the EXIT chart and thus the lower is the error floor. Another approach to minimize error floor is to minimize PEP for perfect a priori knowledge at the demapper. It was shown in [9] that for Rayleigh fading and high SNR this PEP is monotonically decreasing with the harmonic mean $D_h$ of squared Euclidean distances of symbol vectors differing in one bit:

$$D_h = \left( \frac{1}{QM \cdot 2^{QM}} \sum_{\mathbf{x}\in\mathbb{X}} \sum_{n=1}^{QM} \left\| \mu(\mathbf{x}) - \mu(\mathbf{x}+\mathbf{e}_n^{\mathrm{T}}) \right\|^{-2} \right)^{-1}. \quad (9)$$

The first summation is with respect to all $2^{QM}$ permutations $\mathbb{X}$ of $\mathbf{x}$. Addition of $n$th unit row vector $\mathbf{e}_n^{\mathrm{T}}$ corresponds to inversion of the $n$th bit. An optimal MdM for iterative detection should have maximum $D_h$.

## 4.1 MdM with BPSK and QPSK

The following Theorem states optimality for MdMs with BPSK and QPSK for the case of perfect a priori knowledge.

*Theorem 1:* For BPSK and QPSK, i.e., $Q \in \{1,2\}$, and arbitrary number of transmit antennas $M$, $D_h$ is maximized, if a generator matrix of the form

$$G_{QM} = (\mathbf{1}\ \overline{\mathbf{e}_2}\ \overline{\mathbf{e}_3}\ \ldots\ \overline{\mathbf{e}_{QM}}) \quad (10)$$

is applied, followed by two-dimensional mapping $\mu_{2\mathrm{d}}(.)$, as definied in the previous section. The maximum harmonic mean is

$$D_{h,\max} = \frac{4E_\mathrm{s}}{Q} \cdot \frac{(QM)^2}{QM+1} \overset{(2)}{=} \frac{4E_\mathrm{b}\cdot R_\mathrm{c}}{N} \cdot \frac{(QM)^2}{QM+1}. \quad (11)$$

*Proof:* Define the shortest squared Euclidean distance between two distinct symbol vectors as $\alpha$ (see also Fig. 3). For $Q=1$, $\alpha = 4E_\mathrm{s}$, while for $Q=2$, $\alpha = 2E_\mathrm{s}$, thus $\alpha = 4E_\mathrm{s}/Q$. Since each bit of codeword $\mathbf{c}$ is mapped to an independent dimension, it follows that two codewords with a Hamming distance of $l$ bits are mapped to symbol vectors with squared Euclidean distance $l\cdot\alpha$. It is therefore sufficient to maximize the harmonic mean of Hamming distances of codewords pairs $\mathbf{c}$, belonging to input vectors $\mathbf{x}$, which differ in one bit. Define these Hamming distances as $d_{H,i}, i \in \{1,\ldots,QM\cdot 2^{QM}\}$. For the moment, we neglect the facts that codewords have to be distinct and that the $d_{H,i}$ have to be discrete-valued. We only restrict $d_{H,i}$ to be positive and upper bounded by $QM$. Obviously, the harmonic mean is a $\cup$-convex function of the $d_{H,i}$ and is maximal if all $d_{H,i} = QM$. Since this is not allowed for one-to-one correspondence between $\mathbf{x}$ and $\mathbf{c}$ (all entries of $G$ would be 1, thus $G$ would be singular), we decrease as many $d_{H,i}$ as necessary to allow for unique encoding. This would be the case, if all but one inversions of one bit of each of the $2^{QM}$ input vectors $\mathbf{x}$ would yield $d_{H,i} = QM-1$ and one inversion, e.g., of the first bit, would yield $d_{H,i} = QM$. This can be achieved by a linear block code with generator matrix defined in (10).

Hence, we have $2^{QM}$ times $d_{H,i} = QM$ and $2^{QM}\cdot(QM-1)$ times $d_{H,i} = QM-1$. The harmonic mean of these values equals $(QM)^2/(QM+1)$. Multiplication with $\alpha$ yields (11). ■

From the symmetry in (11) with respect to $Q$ and $M$, we can conclude that the error floor of an MdM with $M$ transmit antennas and QPSK is the same as of an MdM with $2M$ transmit antennas and BPSK, if the MdMs are designed according to Theorem 1 and the number of receive antennas is held constant. Note that $G_{QM}$ in (10) is a non-singular matrix with maximum number of 1 entries. Permutations of rows or columns do not change the property of this block code.

## 4.2 MdM with 16-QAM

For 16-QAM Gray mapping $\mu_{2\mathrm{d}}(.)$ the shortest squared Euclidean distance, $\alpha = 2/5\cdot E_\mathrm{s}$, between two distinct symbol vectors can be seen in Fig. 3. However, it is not possible anymore to state that codewords differing in $l$ bits are mapped to symbol vectors with squared Euclidean distance $l\cdot\alpha$. Therefore, Theorem 1 can not be applied straightforward.

We first consider $M=1$. Using a generator matrix $\widetilde{G_4} = (\overline{\mathbf{e}_1}\ \mathbf{1}\ \overline{\mathbf{e}_3}\ \overline{\mathbf{e}_4})$ similar to (10), followed by Gray mapping $\mu_{2\mathrm{d}}(.)$, we find a 16-QAM anti-Gray mapping with $D_h = 10.63\cdot\frac{E_\mathrm{b}\cdot R_\mathrm{c}}{N}$, which is already close to optimum. The following squared Euclidean distances occur: 32 times $5\alpha$, 16 times $8\alpha$ and 16 times $13\alpha$. Only if, in addition, four symbols are swapped, as indicated by the arrows in Fig. 3, or in other words, if we apply $\widetilde{G_4}$ followed by the modified mapping $\mu'_{2\mathrm{d}}(.)$, we obtain the best 16-QAM anti-Gray mapping for an AWGN channel, found by numerical methods in [5]. We denote this mapping as 16-QAM 2d AG. As a consequence of this symbol swapping, half of the symbol pairs, which first had squared Euclidean distances of $8\alpha$ now have $10\alpha$. Computing $D_h$ for this case yields about $10.86\cdot\frac{E_\mathrm{b}\cdot R_\mathrm{c}}{N}$, which is lower than $D_{h,\max}$, if BPSK with $M=4$ or QPSK with $M=2$, respectively, is applied. From (11) we compute $D_{h,\max} = 12.8\cdot\frac{E_\mathrm{b}\cdot R_\mathrm{c}}{N}$. All three schemes transmit $4\cdot R_\mathrm{c}$ information bits per channel use. It was shown in [5] that the 16-QAM 2d AG mapping performs over a Rayleigh channel almost identically as the mapping optimized for Rayleigh fading, which has a slightly higher $D_h = 10.88\cdot\frac{E_\mathrm{b}\cdot R_\mathrm{c}}{N}$.

For $M=2$, we applied all possible 8-dimensional non-singular matrices, where all except seven entries are 1, followed by either $\mu_{2\mathrm{d}}(.)$ or $\mu'_{2\mathrm{d}}(.)$.

The highest value for $D_h$ was obtained with $\widetilde{G}_8 = (\overline{\mathbf{e}_1}\ \mathbf{1}\ \overline{\mathbf{e}_3}\ \overline{\mathbf{e}_4}\ \overline{\mathbf{e}_5}\ \overline{\mathbf{e}_6}\ \overline{\mathbf{e}_7}\ \overline{\mathbf{e}_8})$ in combination with $\mu'_{2\mathrm{d}}(.)$ and is about $25.13 \cdot \frac{E_b \cdot R_c}{N}$, which is again smaller than for BPSK or QPSK at the same information rate: for $Q = 2$ and $M = 4$, we get $D_{h,\max} \approx 28.44 \cdot \frac{E_b \cdot R_c}{N}$. This suboptimality of higher-order modulation ($Q > 2$) is due to the fact that the number $QM$ of bits which are mapped to a symbol vector $\mathbf{s}$, is larger than the number of independent dimensions of $\mathbf{s}$, which is $M$ for real valued and $2M$ for complex valued symbols $s_j$.

Extensive investigation on MdM optimization with the binary switching algorithm [11] was done in [12], but no result could be found, which has a larger $D_h$ than MdM designed with the proposed matrices.

*Theorem 2:* The harmonic mean $D_h$ in (9) is constant with respect to the number of transmit antennas $M$, if conventional symbol-wise mapping is applied.

*Proof:* Compared with a single-antenna transmitter, the frequencies of occurrence of squared Euclidean distances in (9) are $M$ times higher, if $M$ transmit antennas and symbol-wise mapping are applied. This up-scaling does not change the harmonic mean. ■

Hence, if we apply symbol-wise 16-QAM 2d AG for $M = 2$, we have again just $10.86 \cdot \frac{E_b \cdot R_c}{N}$.

## 5 Simulation Results

We have simulated BER as a function of SNR with $10^7$ information bits per SNR value. We use a non-recursive systematic convolutional code with memory 2, code rate $R_c = 1/2$ and generator polynomials (04,07) given in octal numbers, which provides a good match to the demapper in the EXIT-chart, as will be seen later. Interleaving depth, which equals the codeword size, is $S = 96000$. This is rather large in order to factor out effects of short interleaver depths. In all cases we have applied the MdM that provides maximum $D_h$, as explained in Section 4, except in some cases, where conventional symbol-wise mapping is applied, which is indicated by the notation ”2d” in the figures.

Fig. 5 shows BER after 20 iterations (it.) for systems with an information rate of 2 bit/s/Hz. As can be concluded from Theorem 1, the error floors of the $4 \times 2$ BPSK and the $2 \times 2$ QPSK system are the same. Both schemes transmit bits over 4 independent dimensions: 4 spatial dimensions, provided by the i.i.d. channel coefficients, for BPSK and 2 spatial dimensions subdivided into 2 orthogonal dimensions of the complex plane in case of QPSK. This orthogonality among each of the two dimensions provides an advantage for low SNR, visible in an earlier turbo cliff for $2 \times 2$ QPSK. $4 \times 4$ BPSK has an additional receive diversity advantange. BER for 16-QAM 2d AG with $N = 1, 2, 4$ receive antennas, respectively, is also depicted. For $N = 2$ the error floor is higher than that of the two MdMs in Fig. 5 employing the same number of receive antennas, because $D_h$ is smaller as already discussed. For comparison, BER for symbol-wise mappings for $4 \times 4$ BPSK and $2 \times 2$ QPSK with anti-Gray (AG) mapping is shown. No turbo cliff can be observed in the considered SNR interval.

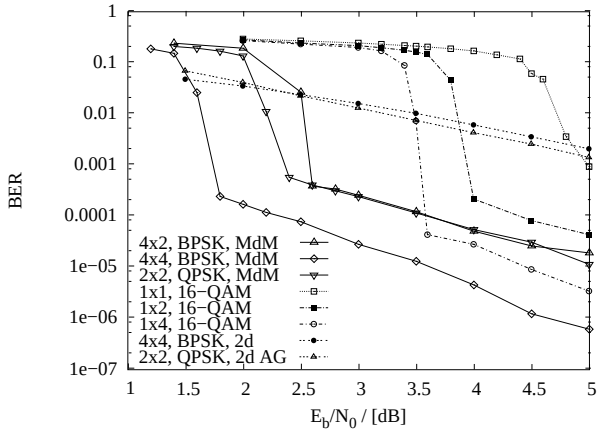


Fig. 5: BER for MdM (—) and conventional symbol-wise mappings (. . . ) for various numbers of tx and rx antennas and modulation orders, 2 bit/s/Hz, $R_c = 1/2$, 20 it.

If we use the previous $M = 4$ BPSK and the $M = 2$ QPSK MdM in a scenario, where the receiver employs only $N = 1$ antenna, we can see from Fig. 6, that their performance is superior to a conventional $1 \times 1$ system with 16-QAM 2d AG mapping. Note that MdM utilizes multiple transmit antennas for spatial multiplexing as opposed to space-time block or trellis codes, where additional redundancy is distributed over these antennas.

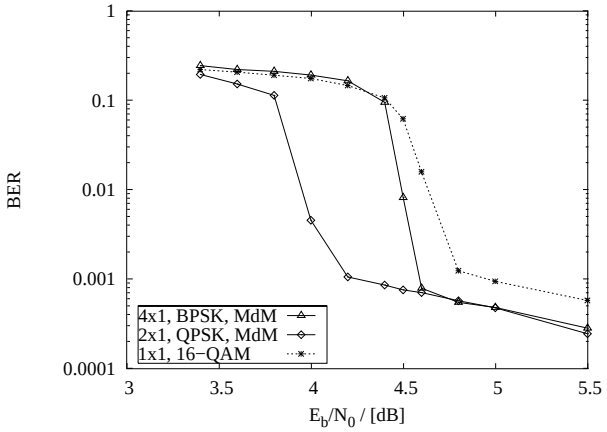


Fig. 6: BER for MdM (—) and conventional symbol-wise mapping (. . . ), $N = 1$ rx antenna, 2 bit/s/Hz, $R_c = 1/2$, 20 it.

In Fig. 7 BER after 40 iterations for systems with information rate 4 bit/s/Hz is depicted. With MdMs no errors could be measured after the turbo cliff. The suboptimality of higher-order modulation discussed in Section 4 can clearly be seen. 16-QAM 2d AG with symbol-wise mapping and $M = N = 2$ is about 3.5 dB worse than the $4 \times 4$ QPSK MdM. Note that from SNR definition (2) an additional 3 dB array gain has to be

considered for the $4\times4$ scheme. BER for the $1\times2$ 16-QAM 2d AG system, which transmits only 2 bit/s/Hz, is also shown. As implied in the discussion of Theorem 2, the error floor is independent of the number of transmit antennas $M$ for symbol-wise mappings. The $M = 4$ QPSK MdM utilizes 8 independent dimensions, while the 16-QAM MdM for $M = 2$ only provides 4. Hence, the turbo cliff of $4\times4$ QPSK MdM occurs earlier.

The optimized $4\times4$ QPSK MdM with outer $R_c = 1/2$ code in Fig. 7 comes very close to the capacity limit. The turbo cliff is at 2 dB, which is only 0.4 dB away from capacity. In [6], $4\times4$ QPSK was combined with an outer turbo-code of rate $R_c = 1/2$ and a turbo cliff at slightly less than 3 dB without measurable error floor was achieved. In [3], a configuration similar to [6] resulted in a turbo cliff at 2.2 dB.

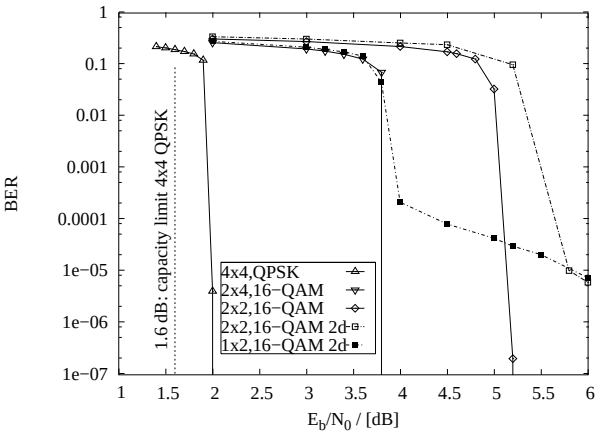


Fig. 7: BER for MdM (—) and conventional symbol-wise mappings (...), 4 bit/s/Hz, $R_c = 1/2$ and $1\times2$ 16-QAM with 2 bit/s/Hz, $R_c = 1/2$, 40 it.

Another advantage of QPSK over 16-QAM MdM becomes visible in the EXIT-chart of Fig. 8. The $I_{E1}(I_{A1})$ curve of $4\times4$ QPSK MdM has an inflection point at about $I_{A1} = 0.5$, which allows a good match to the transfer charactristic of the decoder, resulting in an early turbo cliff. As a drawback, matching of transfer charactristics results in many iterations. A tunnel between both curves is already open at 2 dB. An inflection point of $I_{E1}(I_{A1})$ could not be observed for any 16-QAM MdM. Moreover, $I_{E1}(I_{A1} = 1)$ is smaller in case of the 16-QAM MdM. Symbol-wise anti-Gray QPSK with $M = N = 4$ is also shown.

## 6 Conclusion

In this paper we have presented a method to design MdM by means of generator matrices. This allows optimization with respect to minimum PEP for perfect a priori knowledge at the demapper. For BPSK and QPSK, we proposed optimal generator matrices. QPSK MdMs have an earlier turbo cliff than BPSK MdMs and need just half the number of transmit antennas to achieve the same PEP. Simulation results have shown that MdMs with the same PEP exhibit the same error floor. Our $4\times4$ QPSK MdM allows transmission without measurable error-floor only 0.4 dB away from the capacity limit.

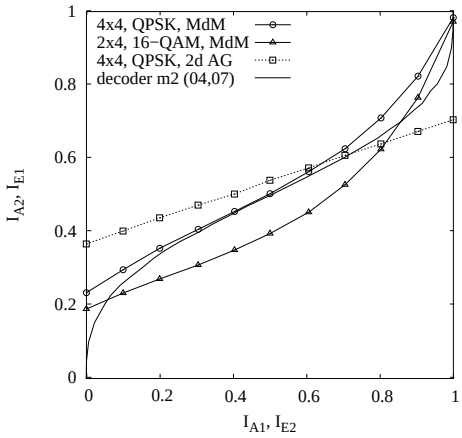


Fig. 8: EXIT-chart at 2 dB, MdM (—) and conventional symbol-wise mapping (...), 4 bit/s/Hz, $R_c = 1/2$

## References


[1] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecom.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[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 *URSI International Symposium on Signals, Systems, and Electronics (ISSSE) 1998*, Pisa, Italy, Sep. 29 - Oct. 2, 1998, pp. 295–300.

[3] S. Bäro, "Turbo detection for MIMO systems: Bit labeling and pre-coding," in *5th International ITG Conference on Source and Channel Coding (SCC) 2004*, Erlangen, Germany, Jan. 14-16, 2004, pp. 11–16.

[4] A. Boronka, N. S. Muhammad, and J. Speidel, "Removing error floor for bit interleaved coded modulation MIMO transmission with iterative detection," in *IEEE Int. Conf. Communications (ICC) 2005*, Seoul, Korea, May 16-20, 2005, pp. 2392–2396.

[5] F. Schreckenbach, N. Görtz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Lett.*, vol. 7, no. 12, pp. 593–595, Dec. 2003.

[6] 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, March 2003.

[7] J. Boutros, N. Gresset, L. Brunel, and M. Fossorier, "Soft-input soft-output lattice sphere decoder for linear channels," in *IEEE Global Telecom. Conf. (GLOBECOM) 2003*, vol. 22, San Francisco, USA, 1. bis 5. Dezember 2003, pp. 1583–1587.

[8] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[9] G. Caire, G. Taricci, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 927–945, May 1998.

[10] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *IEEE Global Telecom. Conf. (GLOBECOM) 1998*, vol. 1, Sydney, Australia, Nov. 8-12 1998, pp. 579–584.

[11] K. Zeger and A. Gersho, "Pseudo-gray coding," *IEEE Trans. Commun.*, vol. 38, no. 12, pp. 2147–2158, Dec. 1990.

[12] F. Breyer, "Untersuchung von multidimensionalen Mappings für MIMO-Systeme," Master's thesis, Institute of Telecommunications, University of Stuttgart, May 2005, number 2116D.