Multilevel Coding for Multiple-Antenna Transmission

Lutz H.-J. Lampe\(^1\), Robert Schober\(^\dagger\), Robert F.H. Fischer

Chair of Information Transmission
University of Erlangen-Nuremberg
Cauerstraße 7/NT, D–91058 Erlangen, Germany
Phone: +49-9131-85-28718, Fax: +49-9131-85-28919, Email: LLampe@LNT.de

\(^\dagger\) Department of Electrical and Computer Engineering
University of Toronto, Canada
Email: rschober@comm.utoronto.ca

Abstract — The application of powerful coding for transmission over multiple-input multiple-output channels is discussed. We emphasize that as an immediate consequence of the mutual information chain rule, multilevel coding (MLC) constitutes the optimum coded modulation scheme. On the other hand, simple bit-interleaved coded modulation (BICM) is only a convenient alternative for the case of two transmit and one receive antennas when combined with orthogonal space-time block codes. Starting from MLC we further propose a hybrid coded modulation scheme, which favorably combines the advantages of MLC and BICM.

\(^1\) Corresponding Author
1 Introduction

Multiple-antenna transmission has attracted considerable attention as means to achieve high bandwidth efficiency and/or high power efficiency over fading channels. Whereas exceptionally high data rates are accomplished by multiple antennas both at the transmitter and at the receiver opening equivalent parallel transmission channels [1, 2], improved reliability of communication is due to increased transmit and/or receive diversity [3, 4, 5, 6, 7].

In this paper, we consider the application of coded modulation to systems with multiple transmit and possibly multiple receive antennas. We concentrate on fast fading channels with channel state information available at the receiver only. Regarding multiple-antenna signaling as multi-dimensional modulation we propose to separate the optimization of coding and modulation. Such an approach has shown to be optimum for single-antenna systems [8]. In perfect analogy, the mutual information chain rule immediately leads to the application of multilevel coding (MLC) with multistage decoding (MSD) [9, 8] in the case of multiple-antenna systems, too.

As modulation strategy we consider both independent signaling over each transmit antenna and space-time block codes (STBC) [10, 11] employing phase-shift keying (PSK) and quadrature-amplitude modulation (QAM), respectively. Here, we exclusively consider STBC’s based on orthogonal designs as studied in [11]. Since orthogonal STBC’s transform the multiple-input multiple-output (MIMO) system into a single-input single-output (SISO) system, cf. e.g. [12, 13], they facilitate the use of coded modulation schemes optimized for single-antenna transmission [14]. In particular, bit-interleaved coded modulation (BICM) [15] is a simple yet power-efficient solution for multiple-antenna transmission with orthogonal STBC’s, cf. [14, 16].

After introducing the considered MIMO and, for the use of STBC’s, the equivalent SISO channel model, MLC for MIMO transmission is described in Section 2. BICM is regarded as alternative coded modulation scheme and also combined with STBC’s. Furthermore, to facilitate practical implementation, a derivative of MLC referred to as hybrid coded modulation (HCM) is proposed. Similar to BICM, HCM relies on Gray labeling of signal points. But in contrast to BICM, HCM is directly applicable to MIMO transmission as will be shown. In Section 3, the different coding
schemes are compared with respect to achievable capacity, and the design of MLC and HCM is illustrated. Simulation results presented in Section 4 are in good accordance with the predictions from information theory. Finally, conclusions are drawn in Section 5.

2 Coding and Modulation for Multiple-Antenna Transmission

We consider the discrete-time system model and complex baseband signaling using $N_T$ transmit antennas and $N_R$ receive antennas.

2.1 Channel Description

At time $k$ ($k \in \mathbb{Z}$ discrete-time index), the vector symbol $x[k] \triangleq [x_0[k] \ldots x_{N_T-1}[k]]^T$ is sent over a MIMO flat fading channel $H[k]$, where $H[k]$ is an $N_R \times N_T$ matrix with entry $h_{\mu\nu}[k]$, $0 \leq \mu \leq N_R - 1$, $0 \leq \nu \leq N_T - 1$, in the $\mu$th row and $\nu$th column accounting for the fading between transmit antenna $\nu$ and receive antenna $\mu$. The symbols $x_{\nu}[k], 0 \leq \nu \leq N_T - 1$, are drawn from an $M$-ary alphabet, e.g., a PSK or QAM constellation, with equal transmit power for each antenna [2]. Collecting the received samples $y_{\mu}[k]$ in a vector $y[k] \triangleq [y_0[k] \ldots y_{N_R-1}[k]]^T$, the input-output relation reads

$$y[k] = H[k]x[k] + n[k],$$

where $n[k] \triangleq [n_0[k] \ldots n_{N_R-1}[k]]^T$ denotes the complex additive white Gaussian noise (AWGN) assumed to be spatially and temporally uncorrelated and with equal variance $\sigma_n^2 = \mathcal{E} \{ \lVert n_{\mu}[k] \rVert^2 \}$, $0 \leq \mu \leq N_R - 1$, at each receive antenna. The fading gains are zero-mean complex Gaussian random processes with variance $\sigma_h^2$ and presumed to be mutually independent for different antenna pairs. We assume fast fading with respect to one coding frame, i.e., an ergodic Rayleigh fading channel is modeled. On the other hand, the fading is slow with respect to one modulation interval and hence, channel state information (CSI) is available at the receiver (but not at the transmitter).

Regarding capacity arguments [2] independent signaling for each transmit antenna is optimum and subsequently applied. For convenience, we choose uniformly, independently, and identically distributed symbols $x_{\nu}[k]$ with variance $1/N_T$, i.e., transmit power is kept constant independent of $N_T$. 
Additionally, for BICM we also consider the use of orthogonal STBC’s [10, 11], i.e., $L \geq N_T$ consecutively transmitted vectors $\mathbf{x}[k]$ constitute one $N_T \times L$ modulation matrix symbol $\mathbf{X}[\kappa]$, $\kappa = [k/L]$, which is addressed by $K$ scalar symbols $x_k[k]$, i.e., the rate is $K/L$ symbols/(channel use). Expecting the channel to remain constant during $L$ channel uses and applying linear processing at the receiver as described in [10, 11], the MIMO fading channel (1) transforms into a SISO fading channel$^2$:

$$y[k] = h[k]x[k] + n[k].$$  \hspace{1cm} (2)

The scalar fading gain $h[k] = L/K \sum_{\nu=0}^{N_T-1} \sum_{\mu=0}^{N_R-1} |h_{\mu\nu}|^2$ is chi-square with $2N_TN_R$ degrees of freedom, the desired signal $x[k]$ has variance $1/N_T$, and $n[k]$ is AWGN with variance $(L/K)N_TN_R\sigma_n^2\sigma_n^2$.

2.2 Multilevel Coding

We assume that the channel realization $\mathbf{H}[k]$ is known to the receiver, and the distribution of $\mathbf{H}[k]$ is known at the transmitter. Then, as we do not aim to optimize the distribution of $\mathbf{x}[k]$, the average mutual information $I(\mathbf{x}; (\mathbf{y}, \mathbf{H}))$ (cf. [2]) constitutes the maximum achievable transmission rate for the MIMO channel. Therefore, subsequently we also refer to $I(\mathbf{x}; (\mathbf{y}, \mathbf{H}))$ as constellation-constrained capacity $C$.

Let $\mathcal{M}(\cdot)$ denote a bijective mapping function of $\ell \triangleq N_T \log_2(M)$ binary digits $b^i, i = 0, \ldots, \ell - 1$, to vectors $\mathbf{x}[k]$ consisting of $N_T$ $M$-ary symbols. Then, by applying the chain rule

$$I(\mathbf{x}; (\mathbf{y}, \mathbf{H})) = \sum_{i=0}^{\ell-1} I(b^i; (\mathbf{y}, \mathbf{H})|b^0, \ldots, b^{i-1}),$$  \hspace{1cm} (3)

the actual vector channel is decomposed into $\ell$ equivalent binary channels or levels. The decomposition (3) immediately suggests that independent encoding of $\ell$ binary data streams at the transmitter and successive decoding at the receiver is optimum. This strategy is known as multilevel coding (MLC) with multistage decoding (MSD) [9, 8]. The separation of coding and modulation via MLC allows us to apply standard binary codes optimized for single-antenna transmission over the AWGN channel. We particularly note that, as well-established for SISO channels, MLC with MSD can approach the constellation constrained capacity $C$ if and only if the code rates $R^i$ are chosen as [8] $^2$For convenience the subscripts $\nu$ and $\mu$ are omitted.
\[ R^i = C_{\text{MLC}}^i \triangleq I(b^i; (y, H)|b^0, \ldots, b^{i-1}), \quad 0 \leq i \leq \ell - 1. \]

In general, a multi-dimensional mapping \( \mathcal{M} \) might be applied, but in terms of capacity any mapping function is optimum. For a practical implementation we follow a pragmatic approach and map blocks of \( \log_2(M) \) bits via a mapping rule \( \mathcal{M}'(\cdot) \) to one constituent PSK/QAM symbol \( x_\nu[k] \) transmitted over one antenna. This means, we sort the levels of MLC first according to the significance of bits and then according to the position \( \nu \) within the vector symbol \( \mathbf{x}[k] \), i.e., bits \([b^0 \ldots b^{\log_2(M)-1}]\) address \( x_0[k] \), bits \([b^{\log_2(M)} \ldots b^{2\log_2(M)-1}]\) address \( x_1[k] \) and so forth. Ungerböck and Gray labeling are commonly used for the constituent mapping \( \mathcal{M}'(\cdot) \) [8]. This implementation of MLC for MIMO transmission is illustrated in Figure 1a).

### 2.3 Bit-Interleaved Coded Modulation

If the dependencies between the levels are neglected for decoding, the mutual information is lowered to

\[ I'(x; (y, H)) \triangleq \sum_{i=0}^{\ell-1} I(b^i; (y, H)) \leq I(x; (y, H)). \]

However, to achieve \( I'(x; (y, H)) \) here it is sufficient to use a single code with rate

\[ R' = \frac{\sum_{i=0}^{\ell-1} I(b^i; (y, H))}{\ell} \]

instead of \( \ell \), in general, different codes. This simple strategy is the well-known bit-interleaved coded modulation (BICM) [15]. Its application to the current situation is depicted in Figure 1b), cf. also [17].

For SISO channels it has been shown [15] that BICM suffers only a negligible loss in capacity compared to MLC/MSD provided that a Gray labeling of signal points is possible. However, it is not obvious whether and how a labeling of vector symbols \( \mathbf{x}[k] \) suitable for BICM can be found.

Since Gray labeling of scalar PSK/QAM symbols is well established and regarding the equivalent SISO channel (2), the combination of BICM with STBC is appealing for MIMO systems, cf. e.g., [14, 16]. Figure 1d) gives the transmitter block diagram for this coded modulation strategy.
2.4 Hybrid Coded Modulation Schemes

The MLC scheme as proposed in Section 2.2 might be disadvantageous for implementation due to the possibly large number $\ell$ of encoders, which also implies shorter code lengths compared to BICM if the transmission delay is limited. To overcome these drawbacks, hybrid coded modulation (HCM) schemes combining the advantages of MLC and BICM are perfectly suited.

In particular, we propose to merge each $\log_2(M)$ levels corresponding to one constituent PSK/QAM symbol as shown in Figure 1c). Thus, the hybrid scheme consists of only $N_T$ levels. MLC and MSD are applied with respect to these $N_T$ levels, but within each level BICM over $\log_2(M)$ bits is performed. The corresponding mutual information follows as ($m \triangleq \log_2(M)$)

$$I''(x; (y, H)) \triangleq \sum_{i=0}^{N_T-1} \sum_{l=0}^{m-1} I(b^{i+m+l}; (y, H)|b^0, \ldots, b^{i+m-1}),$$

(7)

where, in general, $I'(x; (y, H)) \leq I''(x; (y, H)) \leq I(x; (y, H))$ is true. However, applying Gray labeling of constituent symbols, the difference between $I''(x; (y, H))$ and $I(x; (y, H))$ is negligible. Hence, HCM is almost as efficient as MLC in terms of achievable capacity. To approach $I'(x; (y, H))$, the code rates have to be chosen as

$$R^i = 1/m \sum_{l=0}^{m-1} C_{HCM}^{i+m+l} \text{ with } C_{HCM}^{i+m+l} \triangleq I(b^{i+m+l}; (y, H)|b^0, \ldots, b^{i+m-1}), \quad 0 \leq i \leq N_T - 1.$$  

(8)

On the other hand, HCM offers lower implementation complexity and reduced transmission delay compared to MLC.

3 Comparison and Design of Coding Schemes

Based on the constellation constrained capacity the performances of the different coding schemes are compared and a design procedure for MLC and HCM is presented.

3.1 Comparison

In Figure 2 the capacities for the interesting scenarios of $N_T = 2$ (left) and $N_T = 3$ (right) transmit antennas with independent 4PSK signaling are depicted as function of $10 \log_{10}(\tilde{E}_s/N_0)$, where $\tilde{E}_s$ and $N_0$ are the average received energy per symbol at each receive antenna and the one-sided noise.
power spectral density, respectively. We note that by use of Gray labeling for 4PSK symbols, MLC and HCM yield identical performance, i.e., only one curve is shown for each pair \((N_T, N_R)\). For the application of BICM, the capacities for transmission with appropriate STBC’s are also presented. The curves corresponding to a single transmit antenna and BICM serve as reference. In all examples Gray labeling of the constituent constellation is applied.

First, \(N_T = 2\) is considered (Figure 2 left). To support the same maximum rate of 4 bits/(channel use) Alamouti’s STBC \((L = K = 2)\) [10] based on 16QAM symbols and 16QAM with \(N_T = 1\) are employed, respectively. For these schemes BICM performs practically as good as MLC [15].

In case of \(N_R = 1\), the only advantage of multiple transmit antennas is spatial diversity. This diversity is optimally utilized by the STBC. Aiming at target rates of e.g. 2 . . . 3 bits/(channel use), gains of 1 . . . 2 dB in power efficiency are achievable. Clearly, since fast fading offers temporal diversity, which is efficiently exploited by coding, the effect of additional spatial diversity is not as significant as for the block fading scenario, cf. e.g. [1]. Interestingly, independent signaling over two transmit antennas and MLC (HCM) is inferior to space-time block coded transmission. This result is due to the different constituent signal constellations applied.\(^3\) However, the comparison based on different constellations is fair as the number of transmit symbols per modulation interval is equal in all cases and thus, the demodulation complexity becomes comparable. We further observe that BICM with Gray labeling of constituent symbols is not suitable for vector symbols \(\mathbf{x}[k]\), cf. [17, 16].

For \(N_R = 2\), the situation changes\(^4\). Whereas the capability of the MIMO channel is properly utilized by independent transmission for \(N_T = 2\) antennas, the STBC cannot improve over a SISO scheme with increased diversity. Hence, independent 4PSK signaling in combination with MLC (HCM) outperforms the STBC. Remarkably, this is also true for BICM which looses about 0.8 dB compared to MLC (cf. also [17, 16]). The advantage of \(N_T = 2\) over \(N_T = 1\) transmit antennas amounts to about 2 dB for target rates of 2 . . . 3 bits/(channel use).

Next, we discuss the results for \(N_T = 3\) (Figure 2 right). For fixed maximum rate of 6 bits/(channel use)

\(^3\)Of course, regarding unconstrained capacity with Gaussian inputs independent multiple-antenna transmission is optimum.

\(^4\)Note that the signal-to-noise ratio \(E_s/N_0\) is measured at each receive antenna. Accordingly, a gain of 3 dB/4.77 dB for \(N_R = 2/N_R = 3\) over \(N_R = 1\) is simply owing to the increased received energy of the desired signal.
use), the rate 3/4 STBC ($L = 4$, $K = 3$) given in [13] with 256QAM symbols and 64QAM with $N_T = 1$ are employed, respectively\(^5\). Again, BICM is almost optimum for these schemes [15]. As can be seen, in terms of constellation constrained capacity the orthogonal STBC is not a convenient choice for $N_T = 3$. The diversity advantage over transmission with $N_T = 1$ is more than compensated by the rate loss implying the use of the large 256QAM constellation. On the other hand, for the MIMO channel with both three transmit and three receive antennas, considerable gains are achievable with independent signaling and MLC (HCM), e.g. about 5 dB for 4 . . . 5 bits/(channel use). The loss due to BICM is comparatively small for $N_R = 3$, but significant for $N_R = 1$.

### 3.2 Design of MLC and HCM

The application of MLC and HCM requires adjustment of the rates of the $N_T \log_2(M)$ and $N_T$ encoders, respectively. For transmission close to information theoretic limits, the rates have to be chosen according to the capacity of the equivalent channels [8]. Of course, the rate design depends on the labeling of signal points. Whereas HCM requires Gray labeling with respect to the constituent signal constellation, MLC can approach capacity with any labeling. For finite code length however, Ungerboeck labeling has a small advantage [8] and is applied subsequently.

Assuming $N_T = 2$ transmit antennas and 4PSK constellations, Figure 3 illustrates the design for MLC and HCM, respectively. As an example, a rate of 2.5 bits/(channel use) is chosen. The capacities $C_{MLC}^i$ and $C_{HCM}^i$ of the $N_T \log_2(M) = 4$ equivalent channels are plotted. Note that due to Gray labeling of 4PSK, $C_{HCM}^0 = C_{HCM}^1$ and $C_{HCM}^2 = C_{HCM}^3$ holds. In case of MLC, the four address bits are sorted as described in Section 2.2.

Following the capacity rule [8], the target rate 2.5 bits/(channel use) divides optimally into the individual rates $R^i = C^i$ (MLC), $0 \leq i \leq \ell - 1$, and $R^j = 1/m \sum_{l=0}^{m-1} C_{HCM}^{j+m+l}$ (HCM), $0 \leq j \leq N_T - 1$, of the component codes, respectively. The capacities for the particular example are summarized in Table 1. We note that the code rates for $N_R = 1$ are very similar to the respective

\(^5\)Here, it should be noted that the demodulation complexity required for the STBC is increased compared to 4PSK with $N_T = 3$ and 64QAM with $N_T = 1$, respectively.
rates for $N_R = 2$. Hence, a rate design without a priori knowledge about $N_R$ yielding almost optimum performance for both one and two receive antennas is also possible.

4 Simulation Results

The performance of the proposed coding schemes is further illustrated by simulated bit error rates (BER) for $N_T = 2$ transmit antennas and perfectly interleaved, spatially uncorrelated Rayleigh fading channels. In particular, HCM with independent 4PSK signaling and BICM with Alamouti’s STBC and 16QAM symbols are compared for a target rate of 2.5 bits/(channel use). Again, BICM with 16QAM and $N_T = 1$ is used as reference system. In each case, Gray labeling of the constituent constellation is applied. For HCM the rates are assigned according to Table 1. To perform close to the theoretical limits, binary turbo codes [18] (parallel concatenated convolutional codes with 16 states each) with code length 8000 are employed. The code rate is adjusted by symmetric puncturing of parity symbols, cf. [19]. The decoders perform 8 iterations. Both the interleavers of the turbo codes and the bit interleavers are randomly generated.

In Figure 4, the measured BER’s are plotted over $10 \log_{10}(\bar{E}_b/N_0)$ ($\bar{E}_b$: average received energy per information bit at each receive antenna) for $N_R = 1$ (left) and $N_R = 2$ (right), respectively. As reference, using the numerical results from Section 3.1, the corresponding capacity limits taking the finite error rate into account (“rate–distortion capacities”) [20] are also shown. Apparently, the simulated BER’s are in good accordance with the results derived from information theory. For desired bit-error rates of $\text{BER} \leq 10^{-3}$, the coded modulation schemes perform within 1.0 . . . 1.5 dB of the capacity limits. Whereas for one receive antenna BICM with STBC is favorably applied, for two receive antennas HCM with independent transmission from each antenna is superior. With respect to transmission with only one transmit antenna gains of about 1.5 dB ($N_R = 1$) and 2 dB ($N_R = 2$) are achieved, respectively.
5 Conclusions

Powerful coding for transmission over MIMO channels is discussed. In perfect analogy to the SISO channel, the application of MLC as optimum coding scheme for MIMO transmission is an immediate consequence of the mutual information chain rule. In contrast to SISO transmission, BICM is not an appropriate alternative if directly applied to vector symbols, whereas BICM in combination with STBC’s offers high power efficiency for $N_T = 2$ and $N_R = 1$. Starting from the MLC approach, HCM is proposed for multiple-antenna transmission. It is shown that HCM can provide a favorable trade-off between the optimality of MLC and the simplicity of BICM. Hence, HCM is a promising candidate for power- and bandwidth efficient transmission over MIMO channels.

References


Table 1: Capacities of equivalent channels for MLC and HCM. $N_T=2$ and 4PSK. Target rate 2.5 bits/(channel use).

<table>
<thead>
<tr>
<th>$N_R$</th>
<th>MLC</th>
<th>HCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$C^0/C^1/C^2/C^3 = 0.39/0.69/0.59/0.83$</td>
<td>$C^0/C^1/C^2/C^3 = 0.53/0.53/0.72/0.72$</td>
</tr>
<tr>
<td>2</td>
<td>$C^0/C^1/C^2/C^3 = 0.41/0.76/0.50/0.83$</td>
<td>$C^0/C^1/C^2/C^3 = 0.58/0.58/0.67/0.67$</td>
</tr>
</tbody>
</table>
Figure 1: Coding for multiple-antenna transmission. a): MLC. b): BICM. c): HCM. d): BICM with STBC.
Figure 2: Capacity over $\frac{E_s}{N_0}$ for MLC, HCM, BICM, and BICM with STBC over spatially uncorrelated Rayleigh fading channels. Left: $N_T = 2$. Right: $N_T = 3$. Dashed lines: $N_T = 1$ as reference curves.
Figure 3: Capacity $C$ and capacities $C^i$ of the equivalent channels over $E_s/N_0$ for $N_R = 2$ and 4PSK. Left: MLC with Ungerböck labeling. Right: HCM with Gray Labeling. Solid lines: $N_R = 1$. Dashed lines: $N_R = 2$. Dash-dotted lines: Rate design for $C = 2.5$ bits/(channel use).
Figure 4: BER over $\bar{E}_b/N_0$ for transmission with $N_T = 2$ and 2.5 bits/(channel use) over spatially uncorrelated Rayleigh fading channels. $\times$: HCM as designed in Table 1. $\square$: BICM with Alamouti’s STBC. $\circ$: $N_T = 1$ as reference. Dashed lines: respective rate-distortion capacity limits.