Coding for Multiple Antenna Transmission in Fast Fading and in OFDM

Stefan H. Müller–Weinfurtner

Wireless Systems Research Dept., AT&T Labs – Research, 200 Laurel Avenue South, Middletown, NJ 07748, USA

Abstract—**The multiple antenna channel coding problem in Orthogonal Frequency–Division Multiplexing is reconsidered, because with frequency interleaving the effective channel characteristic across tones is rather fast fading and does not comply with the quasi–static channel model widely assumed for space–time codes. We study the fast–fading multiple transmit and receive antenna channel via evaluation of the capacity of the ergodic channel. Capacity comparisons give guidelines on how to jointly adjust coding rate and modulation cardinality. Simulations show that bit–based coding outperforms space–time codes in fast fading and, furthermore, offers larger flexibility in rate adaptation.**

Keywords—**Wireless OFDM, Multiple Antennas, Fast Fading, Channel Coding, Capacity, BICM, Space–Time Coding**

I. INTRODUCTION

CHANNEL coding techniques for wireless communications with transmit and receive diversity is a fairly new field of research [1]. Performance limits in terms of outage capacity are derived in [2] for the quasi–static fading channel, while [3] also covers the capacity for the ergodic fast fading channel.

Space–time codes (STC) [4], [5] are well suited for multiple antenna transmission in quasi–static (i.e., block–fading) environments. Our ultimate aim is to apply codes in Orthogonal Frequency–Division Multiplexing (OFDM) to transmit over multipath channels and we realize that the strong channel variation across tones does not comply with the quasi–static channel model. Adjacent channel coefficients are not independent but together with practical frequency interleaving, the resulting channel can be approximated as independent fast fading.

Obviously, we need to reconsider the multiple antenna channel coding problem in OFDM to make best use of frequency diversity in multipath channels. Recent progress has been made in optimizing the performance of STC in fast fading by searching for better codes [6], [7] or by applying the idea of I–Q (inphase and quadrature component) interleaving to STCs [8] but the question remains, as to whether or not STCs are an appropriate channel coding class for fast(er) fading scenarios.

Based on the following results, we conclude that bit–interleaved coding approaches are able to outperform STCs, because they rely on binary codes instead of being signal–space codes like STCs. This feature leads to larger Hamming distances, which are beneficial in fast fading scenarios. Further, a larger flexibility in rate adaptation is achieved, which is desirable in packet data communication.

II. CHANNEL MODEL

We consider block–wise transmission from n_t transmit antennas (Tx) to n_r receive antennas (Rx). The signal constellation in Tx branch j is $\mathcal{A}_j$ and the independently chosen and equiprobable signal points are transmitted simultaneously. In each frame, L symbols are transmitted per antenna. For modulation interval (time step) ℓ, the transmitted symbols are collected in the hypersymbol ($n_t \times 1$ vector)

$a[\ell] = [a_1[\ell], \ldots, a_{n_t}[\ell]]^\mathsf{T}$. Hence, $a[\ell] \in \mathcal{A}_1 \times \ldots \times \mathcal{A}_{n_t}$, where $\times$ denotes the Cartesian product of constituent signal constellations. The hypersymbol is transmitted via the noisy channel to obtain $r[\ell] = H[\ell]a[\ell] + n[\ell], 0 \leq \ell \leq L - 1$. $H[\ell]$ is the $n_r \times n_t$ matrix with channel coefficients $H_{ij}[\ell]$, which describe the transmission between the j–th transmit and i–th receive antenna. The $n_r \times 1$ vector $r[\ell] = [r_1[\ell], \ldots, r_{n_r}[\ell]]^\mathsf{T}$ represents the received samples. The noise samples in the $n_r \times 1$ noise vector $n[\ell]$ are mutually independent zero–mean complex Gaussian variates so that the covariance matrix is $\mathcal{E}_{n[\ell_1], n[\ell_2]} \{ n[\ell_1]n[\ell_2]^\mathsf{H} \} = \delta[\ell_1 - \ell_2] \sigma_n^2 I$, where $\delta[\cdot]$ is the Kronecker delta, $\sigma_n^2 = N_0/T$ is the variance per complex dimension, and I is the identity matrix. N_0 is the one–sided power spectral density of white noise. We further introduce $\mathcal{E} \{ |a[\ell]|^2 \} = E_a/T$, where E_a is the average energy per hypersymbol, i.e., the average total energy transmitted per time step. Together with $\mathcal{E} \{ |H_{ij}[\ell]|^2 \} = 1, \forall i, j$, and independent $H_{ij}[\ell]$, we have $\mathcal{E} \{ |r[\ell]|^2 \} = n_r E_a/T$ and the average SNR per receive antenna is E_a/N_0.

The following comparisons will be based on finite and discrete signal sets like they are used in practical transmission systems. We introduce the total number of bits carried in one hypersymbol $a[\ell]$ as $\Lambda = \sum_{j=1}^{n_t} \log_2 |\mathcal{A}_j|$, which is equivalent to the number of bitlevels in the transmission scheme. $|\mathcal{A}_j|$ is the cardinality of the constituent signal constellation $\mathcal{A}_j$. We map Λ bits b^λ into one hypersymbol and the binary vector $[b^0, \ldots, b^{\Lambda-1}]$ is the label of the hypersymbol.

III. CODING ARCHITECTURES

A. Space–Time Codes

STCs (space–time codes) are presented in [4], [5]. Binary information enters the encoder, and in each time step a complex–valued symbol per antenna is generated according to a code trellis in such a fashion that diversity and/or coding gain is maximized. All transmit antennas use the signal constellation $\mathcal{A}$. In our view of channel coding, all 2–antenna STC in [4] are effectively rate–$1/2$ codes, so that for each 2×1 output vector $a[\ell] \in \mathcal{A} \times \mathcal{A}$ (which could carry $2 \log_2 |\mathcal{A}|$ bits), only $\log_2 |\mathcal{A}|$ bits enter the STC. The space–time decoder directly operates on the received signal–space samples to estimate the most likely information sequence.

STCs are signal–space codes, which already indicates that they might perform poor in fast fading channels with much time diversity. Furthermore, their inflexible coding rate (in most cases $1/n_t$) is undesirable if one aims to achieve fine–grained link adaptation to make best use of the available channel.

B. Bit–Interleaved Coded Modulation

Bit–interleaved coded modulation (BICM) [9] can be expanded to multiple antenna transmission [3], [10] to obtain ad-

vantages in fast fading channels. A single convolutional code (CC) encodes information bits and the coded bits are cyclically distributed into the Tx branches, where they are bit–interleaved by different interleavers, mapped onto the signal constellation $\mathcal{A}$ used in both Tx branches and transmitted. In the receiver, bit metrics are calculated independently for each bit. The bit metrics are deinterleaved and multiplexed into one stream, which is decoded by a conventional Viterbi algorithm.

The complexity of the bit metric computation is $\sim |\mathcal{A}|^{n_t}$ (exponential in n_t). Via the underlying and well–known convolutional codes, the effective coding rate can be adapted in fine–grained steps by the use of actual rate–k/m CCs or by puncturing of a mother code of rate $1/m$.

IV. LINK–LEVEL CAPACITY

A. Mutual Information

We investigate the conditional mutual information [3]

$$I(\boldsymbol{a}; \boldsymbol{r} \mid \boldsymbol{H}) = H(\boldsymbol{a}) - H(\boldsymbol{a} \mid \boldsymbol{r}, \boldsymbol{H})$$
$$= \Lambda - \mathcal{E}_{\boldsymbol{a}, \boldsymbol{r}} \left\{ \log_2 \frac{\sum_{\tilde{\boldsymbol{a}} \in \mathcal{A}_1 \times \ldots \times \mathcal{A}_{n_t}} p_{\boldsymbol{r}}(\boldsymbol{r} \mid \tilde{\boldsymbol{a}}, \boldsymbol{H})}{p_{\boldsymbol{r}}(\boldsymbol{r} \mid \boldsymbol{a}, \boldsymbol{H})} \right\} \quad (1)$$

(measured in bits per hypersymbol) for one specific known channel realization $\boldsymbol{H}$. We exploited equiprobable transmit hypersymbols $\boldsymbol{a}$, i.e., $\Pr\{\boldsymbol{a}\} = 1/\prod_{j=1}^{n_t} |\mathcal{A}_j| = 2^{-\Lambda}$. For the pdf of $\boldsymbol{r}$, we assume independent white Gaussian noise and have $p_{\boldsymbol{r}}(\boldsymbol{r} \mid \boldsymbol{a}, \boldsymbol{H}) = (\pi \sigma_n^2)^{-n_r} \exp(-|\boldsymbol{r} - \boldsymbol{H}\boldsymbol{a}|^2 / \sigma_n^2)$.

B. Bit Probabilities and Mutual Information on Bitlevels

We obtain the mutual information for the λ–th bitlevel addressing the vector $\boldsymbol{a}[\ell]$ — assuming that all other bit positions in the label of $\boldsymbol{a}[\ell]$ are unknown — as [3]

$$I(b^\lambda; \boldsymbol{r} \mid \boldsymbol{H}) = H(b^\lambda) - H(b^\lambda \mid \boldsymbol{r}, \boldsymbol{H})$$
$$= 1 - \mathcal{E}_{b^\lambda, \boldsymbol{r}} \left\{ \log_2 \frac{\sum_{\tilde{b}^\lambda \in \{0,1\}} p_{\boldsymbol{r}}(\boldsymbol{r} \mid \tilde{b}^\lambda, \boldsymbol{H})}{p_{\boldsymbol{r}}(\boldsymbol{r} \mid b^\lambda, \boldsymbol{H})} \right\}. \quad (2)$$

We assumed equiprobable bit values b^λ, i.e., $\Pr\{b^\lambda\} = 1/2$, and for the conditional pdf of $\boldsymbol{r}$, required in (2), we have $p_{\boldsymbol{r}}(\boldsymbol{r} \mid b^\lambda, \boldsymbol{H}) = |\mathcal{A}(b^\lambda)|^{-1} \sum_{\tilde{\boldsymbol{a}} \in \mathcal{A}(b^\lambda)} p_{\boldsymbol{r}}(\boldsymbol{r} \mid \tilde{\boldsymbol{a}}, \boldsymbol{H})$, where $\mathcal{A}(b^\lambda)$ is that subset of the hypersymbol constellation which complies with the bit values at the respective label positions as demanded by b^λ.

C. Channel with Fast Fading

We assume a fast–fading channel so that $\boldsymbol{H}[\ell]$ is random and takes independent values for each ℓ. The matrix entries $H_{ij}[\ell]$ are mutually uncorrelated zero–mean complex Gaussian random variables with $\mathcal{E}\{|H_{ij}[\ell]|^2\} = 1, \forall i, j$. Hence, the magnitude of each entry follows a Rayleigh distribution. The channel parameter is perfectly known for each ℓ and we are interested in the average mutual information (AMI) for coded modulation (CM). Due to fast fading, the channel can be assumed to be ergodic, so that the desired AMI is obtained by averaging over the given channel statistic. This

yields [9] $I_{\text{CM}} = \mathcal{E}_{\boldsymbol{H}}\{I(\boldsymbol{a}; \boldsymbol{r} \mid \boldsymbol{H})\}$. In perfect correspondence, we introduce the bitlevel AMI of the ergodic channel $I^\lambda = \mathcal{E}_{\boldsymbol{H}}\{I(b^\lambda; \boldsymbol{r} \mid \boldsymbol{H})\}$, which is related to the overall AMI of BICM via [9] $I_{\text{BICM}} = \sum_{\lambda=0}^{\Lambda-1} I^\lambda$. We have $I_{\text{BICM}} \leq I_{\text{CM}}$, because the bitlevels are treated as independent binary channels without exploiting known bits from other levels.

D. Near–Optimum Bit Metrics

For Viterbi channel decoding of the bit–based coding architectures, we use the log–likelihood metric

$$L(b^\lambda \mid \boldsymbol{r}, \boldsymbol{H}) = \ln \frac{\Pr\{b^\lambda = 1 \mid \boldsymbol{r}, \boldsymbol{H}\}}{\Pr\{b^\lambda = 0 \mid \boldsymbol{r}, \boldsymbol{H}\}} \quad (3)$$

$$\approx \ln \frac{\max_{\tilde{\boldsymbol{a}} \in \mathcal{A}(b^\lambda = 1)} p_{\boldsymbol{r}}(\boldsymbol{r} \mid \tilde{\boldsymbol{a}}, \boldsymbol{H})}{\max_{\tilde{\boldsymbol{a}} \in \mathcal{A}(b^\lambda = 0)} p_{\boldsymbol{r}}(\boldsymbol{r} \mid \tilde{\boldsymbol{a}}, \boldsymbol{H})} \quad (4)$$

$$= \frac{1}{\sigma_n^2} \left(\min_{\tilde{\boldsymbol{a}} \in \mathcal{A}(b^\lambda = 0)} |\boldsymbol{r} - \boldsymbol{H}\tilde{\boldsymbol{a}}|^2 - \min_{\tilde{\boldsymbol{a}} \in \mathcal{A}(b^\lambda = 1)} |\boldsymbol{r} - \boldsymbol{H}\tilde{\boldsymbol{a}}|^2 \right), (5)$$

which has reasonable complexity after the log–sum (i.e., nearest representative) approximation. Nonetheless, near–optimum performance is achieved by this simplified metric.

E. Link–Level Capacity Evaluation in Fast Fading

We randomly generate matrix channels and evaluate AMI in a Monte–Carlo integration fashion. The results converge fast and can be considered to be fairly exact. Fig. 1 shows results for AMI in fast fading achievable with transmission schemes for a maximum of $\Lambda = 4$ bits per channel use. This can be done by either transmitting 16QAM from one antenna or by transmitting 4PSK from each of two transmit antennas. For fair comparison, both schemes use either only one receive antenna or optimum receiver processing with two receive antennas.
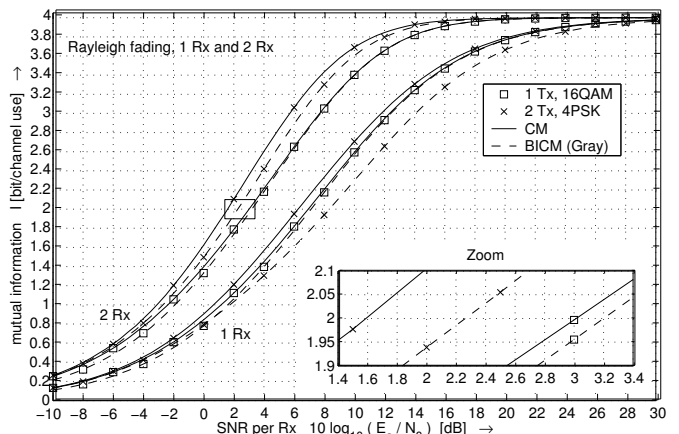

Fig. 1. Mutual information vs. average SNR per Rx antenna for transmission with $\Lambda = 4$ bits per channel use. 16QAM transmitted from one Tx and 4PSK from two Tx are shown.

Let us first consider the case with two receive antennas: We see two curves for each of the transmission schemes. One is for CM (coded modulation) and the other is for BICM (bit–interleaved coded modulation) [9], which strongly relies on the Gray labeling of the (constituent) signal constellations. For 16QAM from 1 Tx, we see the well–known result that the curves for CM and BICM with Gray labeling nearly coincide

for large SNRs, while exhibiting a negligible gap at lower SNRs [9]. To be able to compare BICM to space–time trellis codes, which have effectively rate $1/2$, we look at the horizontal line at spectral efficiency 2. The zoomed area reveals that the SNR loss due to using BICM instead of CM is 0.2 dB for 16QAM. In contrast, for 4PSK from two transmit antennas, the gap between CM and BICM is larger [3] and we find a 0.65 dB loss at spectral efficiency 2. Nonetheless, an SNR advantage of 0.95 dB for the given capacity remains, when moving from 1 Tx 16QAM BICM to 2 Tx 4PSK BICM. Clearly, this is a capacity comparison, only, but BICM actually translates these gains into an appropriately shifted error rate curve as we will see in Section III. Notably, the SNR advantage of multiple transmit antennas over a single transmit antenna is even more obvious at higher spectral efficiencies, so that coding schemes with $R > 1/2$ is an interesting direction to be investigated.

In terms of capacity, one should abstain to operate the 4PSK 2 Tx transmission scheme in fast fading with one Rx antenna, only. It is alarming that the capacity of 2 Tx 4PSK CM is only slightly better than 1 Tx 16QAM CM and for BICM, 2 Tx 4PSK is even worse than 1 Tx 16QAM. It appears that in fast fading, receive diversity is more beneficial than transmit diversity. The reason for this is that the combination of 2 Tx with 1 Rx leads for some channel realizations to ambiguities so that (e.g., for $H_{11} = 1$ and $H_{12} = 1$) the number of distinguishable signal points in the received hypersymbol set can not be guaranteed. A second Rx ensures that this ambiguity occurs much less likely. Maybe this problem is not that visible in capacity computations with Gaussian instead of discrete input signals. Hence, the transmit diversity schemes of order 2 in fast fading should always be operated with $n_r \geq 2$, as otherwise they might be outperformed by higher–order modulation systems with one Tx (i.e., well–defined signal constellation), which have even further advantages in terms of interference cancellation and hence in overall system capacity of cellular systems [11].

The above graph indicates the benefits of transmit diversity to achieve one and the same maximum value of spectral efficiency but does not yet give a design guideline for the combination of coding rate and modulation types to obtain a given overall target spectral efficiency. For this purpose, we compare the AMI of various signal sets for a 2×2 system in Fig. 2.
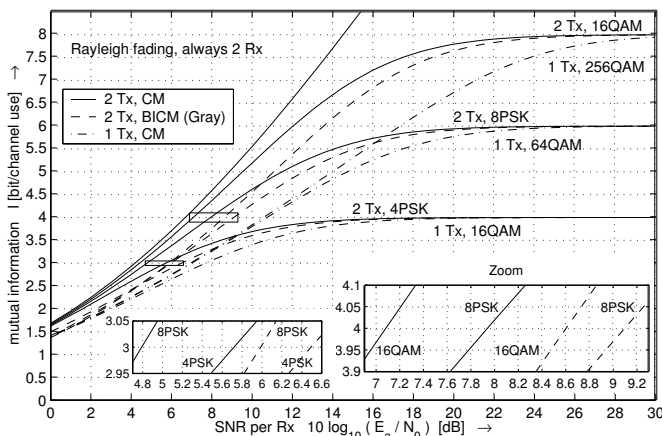
Let us first fix 3 bits per channel use as target in the small zoomed area. Given the larger metric computation complexity as well as higher accuracy requirements for the channel estimation of 8PSK systems when compared to 4PSK, it might be interesting to use 2 Tx 4PSK with rate–$3/4$ code instead of 2 Tx 8PSK with rate–$1/2$ code, especially if BICM is considered, where the capacity gap is 0.5 dB. If we now consider the second zoomed area for 4 bits per channel use, the same reasoning applies to the comparison of 2 Tx 16QAM with rate $1/2$ and 2 Tx 8PSK with rate $2/3$. Here, we have an SNR gap of 0.5 dB for BICM. Note also the large performance gap between BICM and CM, which is 1.5 dB for 16QAM and 1.2 dB for 8PSK. This penalty is lowered if more Rx antennas are used, like in Fig. 3, which shows results for the same transmitter scenario as in Fig. 2 but now with 4 Rx. We again concentrate on 4 bits per channel use and see that in comparison to Fig. 2, the performance penalty for using BICM instead of CM is lowered to 0.65 dB and 0.45 dB for 16QAM and 8PSK, respectively. Hence, a larger number of receive than transmit antennas, i.e., $n_r > n_t$, reduces the capacity gap between BICM and CM. The SNR gap for BICM 16QAM with rate $1/2$ and 8PSK with rate $2/3$ is with 0.55 dB almost the same as the 0.5 dB in Fig. 2.
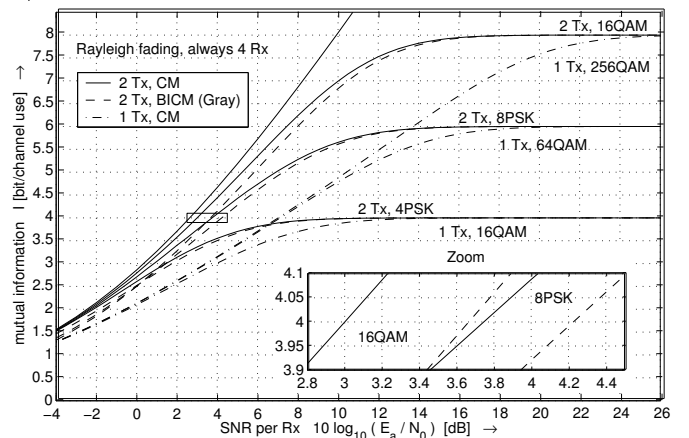
Fig. 3. Mutual information vs. average SNR per Rx antenna (using 4 Rx) for one and two Tx with various maximum spectral efficiencies Λ.

Especially with larger numbers of Tx antennas n_t, the spectral efficiencies with equal signal sets in all Tx antennas increase in steps of n_t, which might be undesirable if the data rate needs to be adjusted in fine–grained steps. The use of mixed signal constellations in the Txs enables a bit–wise adaptation of gross spectral efficiencies during joint design of modulation cardinality and channel coding rate.

V. SIMULATION RESULTS FOR FAST FADING AND OFDM

A. Spectral Efficiency of 2 bit/s/Hz in Fast Fading

In all simulation results, BICM is implemented with sufficient interleaving. For fast fading, the channel matrices are randomly and independently generated. Fig. 4 shows frame error rate (FER) performance vs. average SNR per Rx antenna for systems with 2 Rx and with identical target spectral efficiency of 2 bits per channel use. One frame is $L = 130$ hypersymbols long so that a total of 260 information and termination bits are transmitted. The codes with 16 states use four bits for trellis termination, leaving 256 information bits per frame. We first
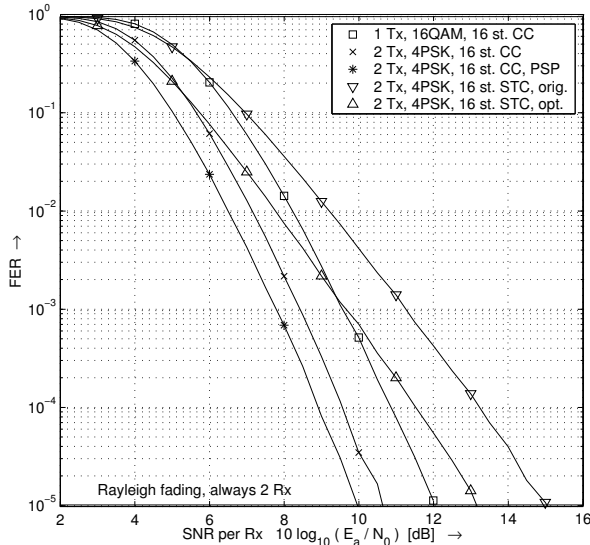
Fig. 2. Mutual information vs. average SNR per Rx antenna (using 2 Rx) for one and two Tx with various maximum spectral efficiencies Λ.

Fig. 4. Frame error rate vs. average SNR per Rx antenna for space–time coding and BICM coding schemes with rate–1/2 convolutional codes and 2 Rx.

consider the two 16 state 4PSK STC (space–time codes). The fast–fading optimized STC [6], [7] (marker △) provides a 1 to 1.8 dB improvement over the original STC [4] (marker ▽), not intended for fast fading. For FER below 10^{-3} (which might not be interesting for some applications) both STC are outperformed by single–Tx 16QAM BICM with rate–1/2 16 state convolutional coding with octal generators $(23, 35)$. But for larger tolerated FER, the optimized STC performs better than the single Tx system. Eq. (5) is used as metric for Viterbi decoding in all BICM simulations. At FER below 10^{-1}, the 2 Tx 4PSK BICM system with rate–1/2 16 state convolutional coding outperforms all others. The slopes of the FER of all BICM schemes are the same, while the slope of the STC systems is significantly less steep, indicating their failure to produce a comparable order of diversity. The BICM schemes enable a higher degree of diversity. It is worth mentioning that the respective spacing of the FER curves of the bit–based systems agrees nicely with the AMI curves in Fig. 1. We actually see the 1 dB shift from 1 Tx 16QAM BICM to 2 Tx 4PSK BICM. Hence, the gains predicted by capacity considerations actually translate into shifted FER, justifying the use of the capacity measure to compare different coding and modulation schemes with multiple antennas.

B. Spectral Efficiency of 4 bit/s/Hz in Fast Fading

We again consider frames of $L = 130$ hypersymbols so that now a total of 520 information and termination bits are transmitted per frame. In Fig. 3, we compared the 2 Tx 8PSK and 16QAM constellations at this spectral efficiency and found a difference of 0.55 dB, only. From results not shown here, we know that the mixed constellation 4PSK/16QAM (with identical average power in the two Tx) shows negligible difference to 8PSK/8PSK in terms of AMI. We want to confirm those similarities by the simulation results in Fig. 5, where we use a rate–1/2 convolutional code with octal generators $(23, 35)$ for 16QAM/16QAM, and a real (non–punctured) rate–2/3 convolutional code with octal generators $(27, 75, 72)$ for 8PSK/8PSK

and 4PSK/16QAM to obtain the spectral efficiency of 4 bits per channel use. Especially for the balanced signal constellations paired with different coding rates, the closeness of the error rate results is striking. The mixed signal constellation performs slightly worse, because the convolutional code has trouble averaging over 3 strongly different bitlevel capacity values I^λ. The balanced transmission schemes only have 2 different capacity values on the bitlevels, so that the convolutional code does not have as much effort to average over them. Expanding the signal constellation in fading 1 Tx/1 Rx links [12] has large benefits which are no longer that visible in the multiple antenna scenario. Good correspondence of simulation and capacity results again justifies the validity of previous capacity comparisons.
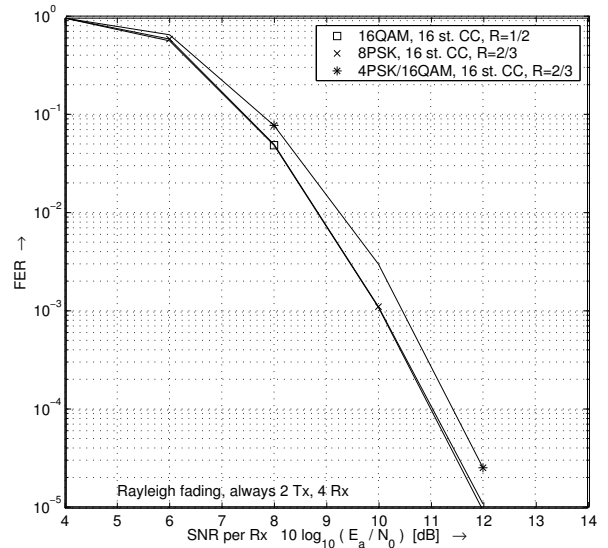

Fig. 5. Frame error rate vs. average SNR per Rx antenna for bit–wrapped coding schemes with rate–1/2 and rate–2/3 convolutional codes for 2 Tx and 4 Rx systems.

C. OFDM Channel Model with Multiple Antennas

The transmission link from Tx antenna j to Rx antenna i is represented by a baud–spaced multipath channel, which is characterized by its finite discrete–time channel impulse response $h_{ij}[\kappa]$, $0 \leq \kappa < K$, $1 \leq i \leq n_r$, $1 \leq j \leq n_t$, i.e., all channels are of length K. The channel taps are zero–mean complex Gaussian random variables and they are mutually uncorrelated in time and also across the antennas. We assume the same average power delay profile for all Tx–Rx links with average tap power $p[\kappa] = \mathcal{E}\left\{|h_{ij}[\kappa]|^2\right\}, \forall i, j$. As indoor model, we use the exponential average power delay profile

$$p[\kappa] = \begin{cases} \dfrac{1 - e^{-1/K_{\exp}}}{1 - e^{-K/K_{\exp}}} \, e^{-\kappa/K_{\exp}}, & 0 \leq \kappa < K \\ 0, & \text{otherwise} \end{cases}, \quad (6)$$

where $K_{\exp}$ is a parameter which characterizes the exponential decay of the average echo power over the channel impulse response length K. From the special scaling, it follows directly that the average sum power is normalized to $\sum_{\kappa=0}^{K-1} p[\kappa] = 1$. The parameter $K_{\exp}$ is loosely related to the rms delay spread.

For OFDM transmission, it is natural to consider coding across the tones, because of delay constraints for data transmission. OFDM has a natural blocking of data so that one

OFDM symbol delay is present in the link. Channel coding over a large number of OFDM symbols would result in mostly unacceptable overall delays. Further, in a sufficiently scattered multipath environment the channel fluctuation in time is usually smaller than the change of channel conditions in frequency, so that a higher degree of diversity results by coding across tones. Hence, the discrete–time variable ℓ from Section II is now associated with the frequency, and via the L–point discrete Fourier transform (DFT), we obtain $H_{ij}[\ell] = \sum_{\kappa=0}^{K-1} h_{ij}[\kappa]\,\mathrm{e}^{-\mathrm{j}2\pi\ell\kappa/L}$. We choose the length of the transmitted signal block L to be a power of 2, in order to implement the transform with a fast transform algorithm. The $n_\mathrm{r} \times n_\mathrm{t}$ channel matrix across the frequency axis ℓ is given by $\boldsymbol{H}[\ell] = [H_{ij}[\ell]]$. With $K < L$ ($K \ll L$), adjacent channel matrices are (strongly) correlated and a frequency–domain interleaving of the transmitted hyper-symbols $\boldsymbol{a}[\ell]$ decreases the error rate.

D. Spectral Efficiency of 2 bit/s/Hz in OFDM

In [13], the use of STC in OFDM is proposed. To have a close relationship to the results in the previous sections, where we investigated block lengths of 130, we choose the DFT size $L = 128$. Further, we use a channel of length $K = 8$ which is fairly short when compared to L and should lead to a conservative estimate, as to whether or not the advantages of bit–based channel coding schemes over STC, which we observed in ideal fast fading, still hold for the OFDM channel properties. As decay parameter, we used $K_\mathrm{exp} = 4.0$. An additional block interleaver of depth 11 for the complex symbols is used for all coding approaches to transform the correlated frequency–selectivity of the OFDM system into a virtually fast(er) fading characteristic. Fig. 6 shows the frame error rate performance of
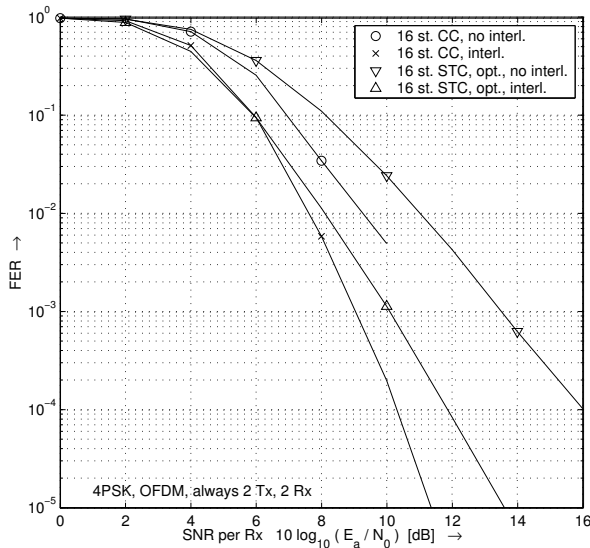

Fig. 6. Frame error rate vs. average SNR per Rx antenna for space–time coding and BICM for 2 Tx and 2 Rx in an OFDM system with $L = 128$ subcarriers and a multipath channel with $K = 8$ and $K_\mathrm{exp} = 4.0$. Results with and without frequency interleaving are shown.

the optimized 4PSK STC [6], [7] and the rate–$1/2$ 4PSK bit–interleaved scheme. Results with and without interleaving are shown for both schemes to demonstrate the necessity of an appropriate frequency–domain interleaving when used in OFDM to destroy the fading correlations between adjacent tones.

Comparing Figs. 4 and 6, we see that the performance of both schemes with interleaving in OFDM is slightly worse than in the ideally fast fading channel, but the performance advantage of the bit–based channel coding is, even though reduced, still remarkable. Further, it is expected that for signal constellations larger than 4PSK the performance advantage of bit–interleaving will be even larger, as observed in Section V-B.

VI. CONCLUSIONS

Binary convolutional coding with appropriate bit–interleaving is a widely accepted way to do channel coding with higher order modulation in fading channels with a single transmit antenna. The results in this paper show that bit–based coding architectures lead to flexible coding schemes for the multiple Tx antenna case, while enabling reasonable performance in fast fading. When a convolutional code exhibits a given distance from the single–Tx capacity limit, it approximately retains this distance from the respective multiple–Tx capacity limit, indicating the usefulness of the bit–based schemes. An interesting possibility lies in higher–rate coding with smaller signal sets, paving the way to combinations like the 2 Tx 8PSK scheme with $R = 2/3$ for a spectral efficiency of 4 bit/s/Hz in Fig. 5.

REFERENCES

[1] G.J. Foschini and M.J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[2] Gerard J. Foschini, "Layered Space–Time Architecture for Wireless Communication in a Fading Environment When Using Multi–Element Antennas," *Bell Labs Technical Journal*, vol. 1, no. 6, pp. 41–59, 1996.

[3] Ezio Biglieri, Giorgio Taricco, and Emanuele Viterbo, "Bit–Interleaved Time–Space Codes for Fading Channels," in *Proceedings of 2000 Conference on Information Sciences and Systems*, Princeton University, Princeton, NJ, USA, 2000, pp. WA4.1–WA4.6.

[4] Vahid Tarokh, Nambi Seshadri, and A.R. Calderbank, "Space–Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.

[5] Ayman F. Naguib, Vahid Tarokh, Nambirajan Seshadri, and A. Robert Calderbank, "A Space–Time Coding Modem for High–Data–Rate Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1459–1478, 1998.

[6] Rick S. Blum, "New analytical tools for designing space-time convolutional codes," in *Proc. of 2000 Conference on Information Sciences and Systems*, Princeton Univ., Princeton, NJ, USA, 2000, pp. WP3.1–WP3.6.

[7] Rick S. Blum, Ye (Geoffrey) Li, Jack H. Winters, and Qing Yan, "Improved space-time coding for MIMO-OFDM wireless communications," *IEEE Trans. on Commun.*, 2001, to appear.

[8] Salam A. Zummo and Saud A. Al-Semari, "Space–Time Coded QPSK for Rapid Fading Channels," in *Proceedings of International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2000)*, London, UK, 2000, pp. 504–508.

[9] Giuseppe Caire, Giorgio Taricco, and Ezio Biglieri, "Bit–Interleaved Coded Modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, 1998.

[10] Ezio Biglieri, Giuseppe Caire, and Giorgio Taricco, "Recent Results on Coding for Multiple–Antenna Transmission Systems," in *Proc. of Int. Symp. on Spread–Spectrum Tech. & Appl.*, NJIT, NJ, USA, 2000, pp. 117–121.

[11] Rick S. Blum and Jack H. Winters, "On the Capacity of Cellular Systems with MIMO," Submitted to ICC, 2002.

[12] U. Hansson and T. Aulin, "Channel symbol expansion diversity," *Electronics Letters*, vol. 31, no. 18, pp. 1545–1546, 1995.

[13] Dakshi Agrawal, Vahid Tarokh, Ayman Naguib, and Nambi Seshadri, "Space–Time Coded OFDM for High Data–Rate Wireless Communications over Wideband Channels," in *Proceedings of IEEE Vehicular Technology Conference (VTC'98)*, Ottawa, Canada, 1998, pp. 2232–2236.