A Supervised Learning Approach to Adaptation in Practical MIMO-OFDM Wireless Systems

Robert C. Daniels, Constantine Caramanis, and Robert W. Heath, Jr.
Department of Electrical and Computer Engineering
The University of Texas at Austin
Wireless Networking and Communication Group
1 University Station, Mail Code: C0806, Austin, TX 78712

***Abstract*—MIMO-OFDM wireless systems require adaptive modulation and coding based on channel state information (CSI) to maximize throughput in changing wireless channels. Traditional adaptive modulation and coding attempts to predict the best rate available by estimating the packet error rate for each modulation and coding scheme (MCS) by using CSI, which has shown to be challenging. This paper considers supervised learning with the k-nearest neighbor (k-NN) algorithm as a new framework for adaptive modulation and coding. Practical k-NN operation is enabled through feature space dimensionality reduction using subcarrier ordering techniques based on post-processing SNR. Simulation results of an IEEE 802.11n draft-compatible physical layer in flat and frequency selective wireless channels shows the k-NN with an ordered subcarrier feature space performs near ideal adaptation under packet error rate constraints.**

I. Introduction

Maximizing the performance of MIMO-OFDM or *multiple-input, multiple-output orthogonal frequency division multiplexing* requires aggressive yet reliable selection of modulation and coding parameters based on channel state information (CSI). In this work, modulation and coding parameters are defined within distinct classes, namely *modulation and coding schemes* (MCS). Selecting the proper MCS based on the observed CSI is referred to as *adaptive modulation and coding* (AMC).

MIMO-OFDM AMC is especially difficult for three primary reasons. First, MIMO complicates the prediction of link performance. The symbol error probability of a MIMO communication signal is not only a function of the signal to noise ratio (SNR), but also other channel parameters [1]. This is exhibited, for example, with the Demmel condition number of the channel matrix as a criterion for switching between diversity and multiplexing operation in MIMO systems [2].

Second, OFDM systems operate within the coherence time of a wireless channel for proper synchronization and detection. The overhead cost in most wireless environments prohibits extensive channel training and physical layer parametrization. Accordingly, some MIMO-OFDM wireless standards, such as IEEE 802.11n, use one set of physical layer parameters for all subcarriers of every OFDM symbol in a physical layer frame. Unfortunately, selecting a single modulation order and coding rate for all subcarriers requires joint performance prediction over all subcarriers simultaneously.

This material is based in part upon work supported by Freescale Semiconductor and the DARPA IT-MANET program, Grant W911NF-07-1-0028.

Finally, most MIMO-OFDM systems include binary convolutional error control codes with OFDM symbol bit interleaving [3]. All subcarriers within an OFDM symbol do not necessarily experience the same fading realizations. On the other hand, the channel is typically constant throughout the duration of a packet. Coding is performed over all bits together in a packet, hence neither fast fading nor flat fading channel distributions accurately model the channel observed.

As a consequence of the aforementioned complexities, MIMO-OFDM AMC is not straightforward. Past efforts have attempted to predict packet error rate based on bit-interleaved convolutional coding performance approximations [4] or by computing look-up-tables for various link quality metrics [5]. In general, previous efforts either suffer from inaccuracy due to model-based approximation or lack implementation clarity due to large-dimension look up tables. A MIMO-OFDM AMC methodology that enables clear implementation and is able to account for the numerous system complexities is highly desired.

In this paper we propose a paradigm shift in the treatment of the MIMO-OFDM AMC problem by performing AMC within a supervised learning framework. Every wireless channel realization (for each packet) is classified or matched to a distinct MCS through the use of a feature set extracted from the CSI. We will treat the system as a black box and use observations (features) associated with past system performance (i.e. the *training set*) to predict the performance output of the black box given a new set of feature values.

A supervised learning approach to AMC is attractive for several reasons including flexibility in the feature set and the ability for the training set of each classifier to be actively updated as more realizations of the channel are observed during typical network use. Performance related data that is often discarded, for example cyclic redundancy checks along with the associated MCS, can be harnessed to update the training set online and actively improve AMC performance during the lifetime of the wireless network. Moreover, the black box approach to the MIMO-OFDM system allows for learning techniques to apply adaptation even when linear system models break down due to non-ideal operation of analog and RF circuit components. Applications of machine learning to adaptation in wireless networks have been suggested in the past [6], [7], but not with the detail or in the MIMO-OFDM framework considered here.

II. System Model

This section defines the MIMO-OFDM system model under investigation for AMC in Section II-A. The application of k-Nearest Neighbor to packet-based MIMO-OFDM AMC is defined in Section II-B.

A. MIMO-OFDM Model

For a given OFDM symbol, $m \in \{0, 1, \ldots, N_{\text{OFDM}} - 1\}$, and subcarrier $n \in \{0, 1, \ldots, N-1\}$, the linear time invariant frequency domain impulse response relationship between the input symbols, $\mathbf{X}[m,n]$, and the output symbols, $\mathbf{Y}[m,n]$, is defined as

$$\mathbf{Y}[m,n] = \sqrt{E_s}\mathbf{H}[n]\mathbf{X}[m,n] + \mathbf{V}[m,n] \tag{1}$$

where N is the DFT size, $\mathbf{X}[m,n] \in \mathbb{C}^{N_t}$, $\mathbf{V}[m,n] \in \mathbb{C}^{N_r}$, $\mathbf{Y}[m,n] \in \mathbb{C}^{N_r}$, and $\mathbf{H}[n] \in \mathbb{C}^{N_r \times N_t}$ given N_t, N_r represent the number of transmit, receive antennas respectively. Assuming $\mathbb{E}[|\mathbf{X}[m,n]|^2] = 1 \; \forall \; m, n$, E_s designates the expected total transmitted signal energy. We will consider complex Gaussian noise vectors $\mathbf{V}[m,n]$ where each element has variance σ^2 and mean 0. The wireless channel $\mathbf{H}[n]$ observes the quasi-static block fading assumption, such that the channel stays consistent for all OFDM symbols within a packet. For the remainder of this paper we will assume all OFDM symbol subcarriers are a member of $\mathbb{C}^{N_t}_{M\text{QAM}}$, the QAM constellation of order M. AMC selects the appropriate QAM constellation order M for symbols $\{\mathbf{X}[m,n]\}_{n=0}^{N-1}$, the convolutional coding rate for the bits mapped into the QAM symbols, as well as the number of spatial streams in $\{\mathbf{X}[m,n]\}_{n=0}^{N-1}$ for the duration of the physical layer frame. MIMO communication theory indicates the potential number of spatial streams $N_s \leq \min\{N_r, N_t\}$.

For our study we will follow the MCS specifications from IEEE 802.11n. Each class $i \in \{0, 1, \ldots, I-1\}$ for MCS_i defines a convolutional coding rate, M, and N_s to determine the physical layer throughput $\mathcal{R}_i$. This AMC framework can be defined explicitly as a classification problem using the channel state information, the transmitted signal energy, and receiver noise variance.

AMC Classification: Using a feature set extracted from $\{\mathbf{H}[n]\}_{n=0}^{N-1}$, E_s, and σ^2 we select class i (corresponding to MCS_i) to optimize the performance metric. This paper shall consider the packet error rate performance metric.

- A class i is only selected if the corresponding packet error probability of MCS_i is less than or equal to $\mathcal{P}$ (i.e. $\text{PER}_i \leq \mathcal{P}$). Therefore, classification with this performance metric uses class

$$\arg\max_i \{\mathcal{R}_i : \text{PER}_i < \mathcal{P}\} \tag{2}$$

for a given channel realization.

Packet error rate was chosen as a performance metric since it is widely observed in past AMC work [5]. It should be noted that eq. (2) will not optimize the rate over the entire distribution of the channel given PER constraint $\mathcal{P}$. Without prior knowledge of the channel distribution, we ensure each channel realization meets the packet error rate constraint.

B. k-Nearest Neighbor AMC

k-NN is a supervised learning algorithm to select a class $i \in \{0, 1, \ldots, I-1\}$ based on a set of training data. Each element in the training set, indexed by $w \in \{0, 1, \ldots, W-1\}$, is known to belong to a distinct class i. Implementation of the algorithm on a real data set uses a distance metric $d(\bullet)$ to determine the "neighbors" in the training set. Therefore, given a query from a channel observation we will search through the training set and find the k nearest neighbors that minimize the distance metric. The class assigned to the query is the class that occurs most often among the k-nearest neighbors in the training set [8].

Application of k-nearest neighbor to the AMC classification problem requires some simple modifications. The expected performance of a MIMO-OFDM system is uniquely determined by $\{\mathbf{H}[n]\}_{n=0}^{N-1}$, E_s, and σ^2. However, each member of the training set and all queries are represented by a p-dimensional vector in $\mathbb{R}^p$. In the k-NN AMC classification problem, a transformation P maps the observations to the feature space, i.e. $P(\{\mathbf{H}[n]\}_{n=0}^{N-1}, E_s/\sigma^2) = \mathbf{z}$ (or $\mathbf{q}$) to reduce the feature dimension and improve performance. The feature vector $\mathbf{z}$ and query $\mathbf{q}$ will be associated with a distinct class $i \in \{0, 1, \ldots, I-1\}$ corresponding to MCS_i.

Training the k-NN system is completed with W distinct realizations of $(\{\mathbf{H}[n]\}_{n=0}^{N-1}, E_s/\sigma^2)$ in the training set. Each distinct realization $w \in \{0, 1, \ldots, W-1\}$ is assigned to a class i according to (2) and $\mathbf{z}_w$ indicates the feature vector corresponding to realization w. In total, training consists of a mapping between the feature set for all realizations and its corresponding modulation and coding scheme according to (2)

$$\{\mathbf{z}_w\}_{w=0}^{W-1} \mapsto \{i(w)\}_{w=0}^{W-1} \tag{3}$$

where $i(w) \in \{0, 1, \ldots, I-1\}$ is the ideal AMC classification. Given the training feature set $\{\mathbf{z}_w\}_{w=0}^{W-1}$ and its associated classes $\{i(w)\}_{w=0}^{W-1}$, $k-$NN uses the following algorithm. We will assume that we are given a query $\mathbf{q} \in \mathbb{R}^p$ in the feature space and arbitrary k with distance metric $d(\mathbf{z}_w, \mathbf{q})$.

Algorithm *k-NN for AMC*
1. $n_i \leftarrow 0 \; \forall \; i \in \{0, 1, \ldots, I-1\}$
2. **for** $a \leftarrow 1$ **to** k
3. $\quad w_a \leftarrow \arg\min_w \{d(\mathbf{z}_w, \mathbf{q}) : w \notin \{w_1, \ldots, w_{a-1}\}\}$
4. $\quad n_{i(w_a)} \leftarrow n_{i(w_a)} + 1$
5. modeset $= \arg\max_i \{n_i\}$
6. **return** $\min\{\arg\min_i \{\mathcal{R}_i : i \in \text{modeset}\}\}$

One notable feature of this algorithm is the rule for breaking ties. If two distinct classes $i' \neq i''$ return the same number of neighbors to the query, the tie is broken using the rate. That is, without loss of generality, i' is selected if $\mathcal{R}_{i'} < \mathcal{R}_{i''}$. It is possible for higher rate MCSs to be more reliable than a lower rate MCS, but this only occurs during atypical frequency-selective fading realizations [5]. $\mathcal{R}_{i'} = \mathcal{R}_{i''}$, then we will

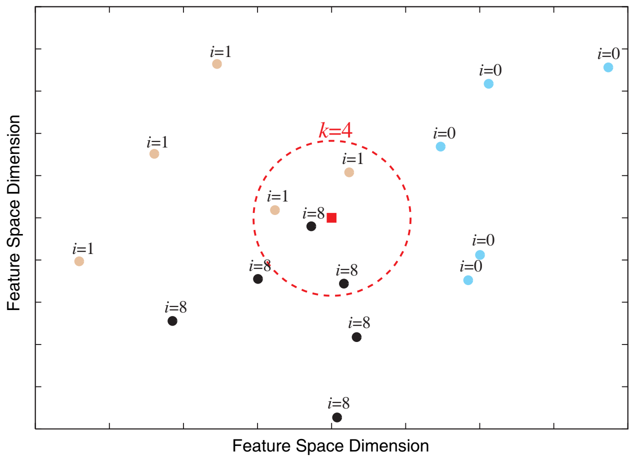


Fig. 1. 4-NN AMC that assigns MCS_1 since $\mathcal{R}_8 = \mathcal{R}_1$ and $1 < 8$

TABLE I
SIMULATION PARAMETERS FOR k-NN AMC IN IEEE 802.11N

Property	Value
SNR	$10\log_{10}(E_s/\sigma^2) \in \{1, 2, \ldots, 28\}$
Classes (i)	$i \in \{0, 1, \ldots, 15\}$
PER Target ($\mathcal{P}$)	0.1 (10% PER)
Realizations (W)	$1000 \times 28 = 2.8 \times 10^4$
$N_r \times N_t$	$2 \times 2 \Rightarrow N_s \in \{1, 2\}$
k	25
$d(\bullet)$	Euclidean distance

always select the lower index, i.e. i' is selected since $i' < i''$. The lower index is chosen because we have assumed that two classes with the same rate have a different number of spatial streams. We assume the lower class index will always use a lower number of spatial streams, as is the case in IEEE 802.11n. This conservative rate/spatial stream classification suggests that the MCS selection is biased towards lower rate selection and a lower number of spatial streams. Hence, we are deciding to choose reliability over rate. This operation is summarized in Figure 1.

III. SUPERVISED LEARNING FOR AMC IN IEEE 802.11N

In this section we apply our k-NN AMC algorithm to the physical layer of an IEEE 802.11n system. Section III-A provides an exposition of the system parameters as well as implicit assumptions in k-NN operation. Section III-B discusses the feature set selection used for the simulation results in Section III-C.

A. System Assumptions

The basic parameters of this IEEE 802.11n-based study of k-NN AMC are summarized in Table I. We observe many system assumptions that will impact the results:

- *Fixed Packet Length*: All packets are 128 Bytes in length. Packet length, in general, impacts adaptation performance since packet error rate is a function of the number of bits per packet as well as the bit error rate.
- *Perfect Synchronization*: We assume that our device is perfectly synchronized in time and frequency. In practice, training symbols and pilot tones are used to synchronize the packet. We will neglect the pilot tones and training sequence during our study since our primary interest is a preliminary study of supervised learning for MIMO-OFDM AMC.
- *Perfect Channel Knowledge*: Although the receiver does estimate the wireless channel for equalization, we assume that k-NN AMC knows the effective channel at the receiver perfectly. This allows us to eliminate k-NN performance degradation effects due to channel estimation inaccuracy.
- *Linear Equalization*: Prediction of wireless system performance is significantly impacted by the specific receiver implemented. Equalization in this paper uses a soft-output zero-forcing structure in the frequency domain to compute log likelihood ratios for a Viterbi detector.
- *No Space-Time Processing*: Other than cyclic delay diversity no space-time processing techniques such as space-time block coding or precoding occur during simulation.

B. Feature Space Selection and Subcarrier Ordering

As a first step we will consider communication theory results to advise selection of the feature space. Since we are operating a 2×2 system and selecting MCSs that use 1 or 2 spatial streams (MCS $0-15$ in IEEE 802.11n)[1], we must determine the performance in both scenarios. The performance of a frequency-flat uncoded single carrier MIMO link with a fixed effective channel matrix $\tilde{\mathbf{H}} \in \mathbb{C}^{N_r \times N_s}$, signal energy E_s, receive antenna noise variance σ^2, and linear zero-forcing (ZF) equalizer $\mathbf{G}_{\text{ZF}} = \tilde{\mathbf{H}}^\dagger \in \mathbb{C}^{N_s \times N_r}$ is a function of the post-processing SNR [9][2]. Post-processing SNR can be defined for any linear equalizer. The ZF post-processing SNR for spatial stream $a \in \{1, 2, \ldots, N_s\}$, in general, is defined as $\text{SNR}_a^{\text{ZF}} \triangleq E_s/(N_s\sigma^2[\tilde{\mathbf{H}}^*\tilde{\mathbf{H}}]_{a,a}^{-1})$. This result demonstrates that the uncoded BER of subcarrier n in a MIMO-OFDM system is a function of $\mathbf{H}[n]$, E_s, and σ^2. With coding, however, the performance is much more complicated for all but very simple channel scenarios.

One simple scenario, the frequency flat channel (single tap), allows all subcarriers to be treated equally. Therefore, performance prediction is available as a function of SNR_a^{ZF} for any stream a [10]. To demonstrate this, the first two plots in Figure 2 visualize the k-NN AMC classification of MCS $0-7$ and MCS $8-15$, respectively with $[\mathbf{H}[n]]_{\ell,m} \sim \mathcal{CN}(0,1)$ for $\ell \in \{1, 2, \ldots N_r\}, m \in \{1, 2, \ldots N_t\}$, $\mathbf{H}[n] = \mathbf{H}[n']$, and $n, n' \in \{0, 1, \ldots, N-1\}$. Classification is performed using post-processing SNR averaged over all subcarriers. There are

[1]MCS $0-7$ use 1 spatial stream and define rates from 6.5 Mbps (MCS 0) to 65.0 Mbps (MCS 7). MCS $8-15$ use 2 spatial streams and define rates from 13.0 Mbps (MCS 8) to 130.0 Mbps (MCS 15).

[2]$[\mathbf{A}]_{a,b}$ denotes the entry in row a and column b of matrix $\mathbf{A}$. $\mathbf{A}^\dagger$ denotes the Moore-Penrose pseudoinverse of matrix $\mathbf{A}$. Effective channel $\tilde{\mathbf{H}}$ is a combination of the wireless channel and linear transmit processing (diversity combining).

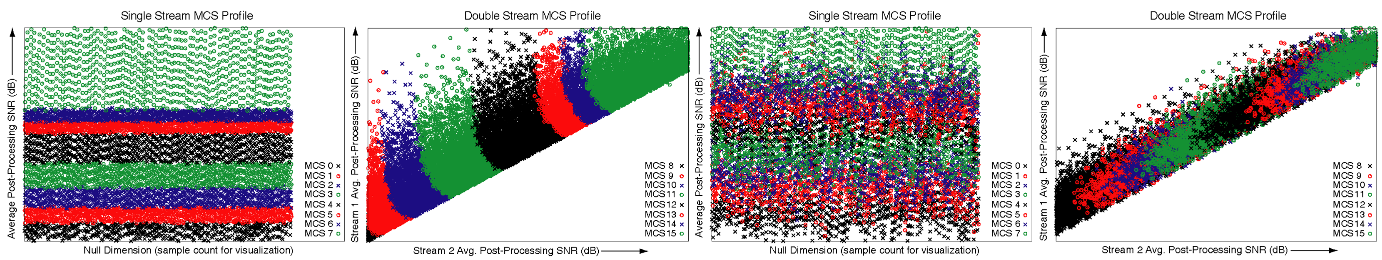


Fig. 2. Single and double stream MCS visualization for flat and frequency selective channels using average SNR feature space

three important notations for this figure. First, instead of classifying MCS $0-15$ jointly, we have chosen to classify MCS $0-7$ and MCS $8-15$ separately. In general, MIMO systems require different performance metrics since the effective channel changes as N_s changes. Therefore, rather than performing k-NN on one feature set with mixed performance identifiers associated with different numbers of spatial streams, we perform k-NN individually over all MCS sets with the same performance identifiers (i.e. the same number of spatial streams). Once we obtain the suggested MCS for each spatial stream, final MCS selection is performed by picking the MCS with highest rate. Remember that k-NN classification must satisfy the PER constraint, so any chosen MCS, regardless of the number of spatial streams, will meet the PER constraint. Second, the abscissa on the single stream classification is only used for visualization and does not represent a feature space dimension. Finally, note that one dimensional post-processing SNR for MCS $0-7$ and the two-dimensional post-processing SNR for MCS $8-15$ successfully separate each modulation and coding scheme. Stream 1 represents the stronger post-processing SNR and Stream 2 represents the weaker post-processing SNR.

For frequency flat channels we successfully implemented k-NN AMC in IEEE 802.11n using a 1-dimensional feature set ($p=1$) and 2-dimensional feature set ($p=2$) for classification of MCS $0-7$ and MCS $8-15$, respectively. Unfortunately, modern systems operate over large spectrum allocations with frequency selective fading. Moreover, cyclic delay diversity is often implemented to prevent destructive combining and add frequency diversity that is captured over all subcarriers through convolutional codes [11]. As a result, post-processing SNR averaged over all subcarriers does not uniquely determine the performance of a system. The third and fourth subfigures in Figure 2 show the classification of MCS $0-7$ and MCS $8-15$, respectively, in a three-tap channel with uniform power delay profile and cyclic delay diversity using average post-processing SNR feature sets. Plainly, these features no longer provide well separated classes and are insufficient for k-NN AMC.

Since the performance of the system is uniquely determined by the channel matrix for all subcarriers as well as the signal energy and noise variance, the feature space of the training set might include $(\{\mathbf{H}[n]\}_{n=0}^{N-1}, E_s/\sigma^2)$. Unfortunately, due to the curse of dimensionality for supervised learning, dimensions of order $N_r \times N_t \times 2 \times N$ for typical values of N_t, N_r, and N require exceedingly large training sets [12]. Practically speaking, the number of dimensions feasible for the feature space depends on $\mathcal{W}$ and classification algorithm complexity constraints.

To reduce the feature space dimension we exploit the existence of a single QAM modulation order across all subcarriers. Therefore, the probability of symbol error in each subcarrier fluctuates drastically within an OFDM symbol of frequency-selective channels. Assuming subcarrier position does not affect total packet performance (which is not always true due to imperfect bit interleaving), it seems plausible that analysis of the per-subcarrier post-processing SNR distribution without respect to subcarrier location will determine packet performance. Let $\text{SNR}_a^{\text{ZF}}[n]$ be the post-processing SNR for spatial stream a of data subcarrier $n \in \{0, 1, \ldots, N_{ds}-1\}$. In IEEE 802.11n the number of data subcarriers (DFT size minus pilot subcarriers and guard tones) $N_{ds} = 52$ for 20 MHz channels. We are interested in observing the SNR subcarrier distribution for different channel realizations. Accordingly we define $\text{SNR}_{(\eta)}^{\text{ZF}}$ to be the ηth smallest element in the set

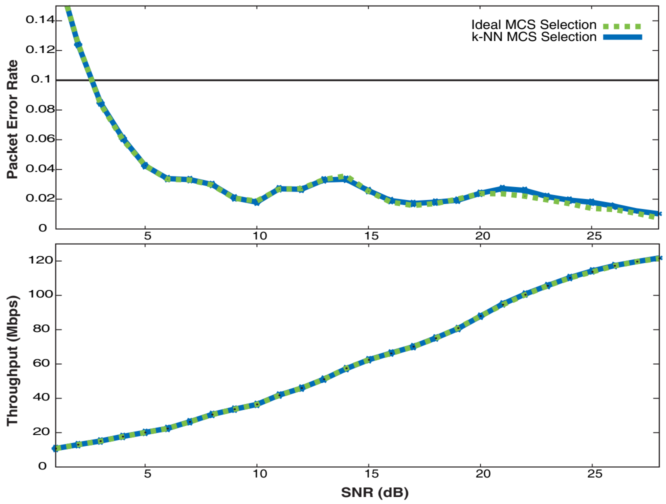
$$\left\{ \left\{ \text{SNR}_a^{\text{ZF}}[n] \right\}_{a=1}^{N_s} \right\}_{n=0}^{N-1}. \tag{4}$$

This method, namely subcarrier ordering, was motivated by empirical observations in IEEE 802.11n channels that suggested $\text{SNR}_{(\eta)}^{\text{ZF}}$ for a few values of $\eta \in \{1, 2, \ldots, N_{ds}N\}$ often determined whether a packet was correctly received.

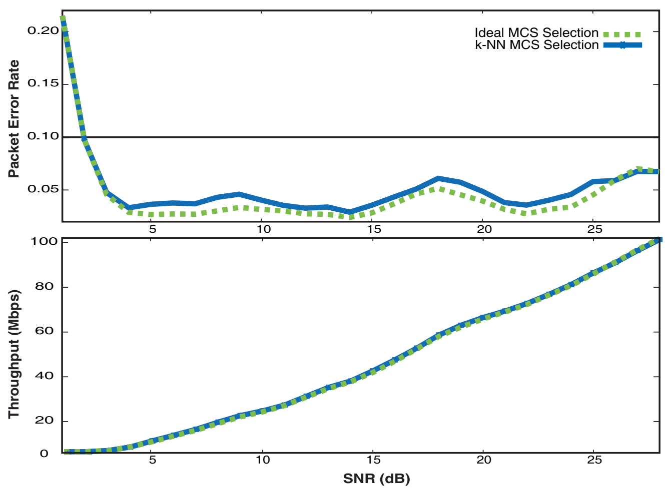
Extensive computer searches have been utilized to discover appropriate feature spaces. Experimentally we have determined that only 4 dimensions are necessary for successful k-NN AMC of MCS $0-7$ and MCS $8-15$ in IEEE 802.11n systems. These computer searches suggest a training set with good performance consists of features

$$\mathbf{z} = [\text{SNR}_{(1)}^{\text{ZF}}, \text{SNR}_{(2)}^{\text{ZF}}, \text{SNR}_{(6)}^{\text{ZF}}, \text{SNR}_{(26)}^{\text{ZF}}]^T \tag{5}$$

for each channel realization. MCS $0-7$ determines these ordered subcarrier feature sets by examining only a single spatial stream, while MCS $8-15$ defines the features through ordering over both spatial streams. These feature sets were discovered to produce the least number of classification errors by searching through all possible ordered subcarrier selections in fixed dimensions. Intuitively, the most important performance considerations are the worst case subcarriers and the average subcarrier. The subcarriers with poor SNR often dictate what



(a) Frequency Flat Channel Performance



(b) Frequency Selective Channel Performance

Fig. 3. Average throughput and PER of k-NN AMC for IEEE 802.11n systems using per-stream classification and subcarrier ordering based feature sets

modulation order the system uses, since they are most likely to yield bit errors when all subcarriers use the same modulation.

C. Performance Results

After establishing the training set with 10% PER constraints, Figure 3(a) shows the average throughput and packet error rate as a function of SNR for frequency flat channels. As expected, the k-NN AMC algorithm with an average post-processing SNR feature space provides nearly ideal performance. Figure 3(b) shows the same performance results using the 4 dimension ordered subcarrier feature space for the frequency selective channels and cyclic delay diversity. Remarkably, there is almost perfect performance in this scenario as well. Close inspection of the performance of k-NN in frequency selective channels shows that we selected the wrong class about 10% of the time. However, all of these errors occurred on boundary regions between classes in the feature space. Therefore, sometimes we selected an MCS with a PER just above $\mathcal{P}$ and sometimes we selected an MCS with a conservative PER. Over time these errors averaged out, yielding negligible performance degradation. This reflects our consideration of the PER threshold as a "soft" constraint. As a final note, k was set equal to 25 in order to average out errors in the training set. Even though the training set is performed offline with perfect knowledge of the channel, determination of the packet error rate given a channel and an associated SNR is subject to estimation error. Moreover, large k allows us to reduce the effect of the conservative tie breaking method included in our k-NN AMC algorithm. If we select k too large, however, we will suffer from excessive processing complexity and will eventually start adding errors due to limited data set sizes.

IV. Conclusions

Supervised learning with the k-nearest neighbor algorithm has emerged as an approach for adaptive modulation and coding in modern MIMO-OFDM systems. We were able to demonstrate exceptional performance is frequency selective channels (uniform power delay profile). This was enabled through subcarrier ordering for dimensionality reduction. Perhaps the most valuable observation is that this algorithm is easily extended to different channel models without retraining the system. Since our choice of the training set was very general and because our feature set does not entirely depend on the channel distribution, this approach is very promising for real systems.

References

[1] S. Catreux, V. Erceg, D. Gesbert, and R.W. Heath, Jr., "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Communications Magazine*, vol. 40, no. 6, pp. 108–115, Jun 2002.

[2] R.W. Heath, Jr. and A.J. Paulraj, "Switching between diversity and multiplexing in MIMO systems," *IEEE Transactions on Communications*, vol. 53, no. 6, pp. 962–968, June 2005.

[3] M.R. McKay and I.B. Collings, "Capacity and performance of MIMO-BICM with zero-forcing receivers," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 74–83, Jan. 2005.

[4] Fei Peng, Jinyun Zhang, and W. E. Ryan, "Adaptive Modulation and Coding for IEEE 802.11n," in *Proceedings of IEEE Wireless Communications and Networking Conference*, Kowloon, Mar.11–15, 2007, pp. 656–661.

[5] S. Kant and T. L. Jensen, "Fast link adaptation for IEEE 802.11n," M.S. thesis, Aalborg University, February 2007.

[6] Yunqian Ma, "Improving wireless link delivery ratio classification with packet SNR," in *Proceedings of the IEEE International Conference on Electro Information Technology*, May22–25, 2005.

[7] Gang Xu and Yourong Lu, "Channel and Modulation Selection Based on Support Vector Machines for Cognitive Radio," in *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing*, Wuhan, Sept.22–24, 2006, pp. 1–4.

[8] Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification (2nd Edition)*, Wiley-Interscience, November 2000.

[9] R.W. Heath, Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Communications Letters*, vol. 5, no. 4, pp. 142–144, Apr 2001.

[10] Shu Lin and Daniel J. Costello, *Error Control Coding, Second Edition*, Prentice Hall, April 2004.

[11] G. Bauch, "Capacity optimization of cyclic delay diversity," in *Proceedings of the IEEE Vehicular Technology Conference*, 2004, vol. 3, pp. 1820–1824 Vol. 3.

[12] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning*, Springer, August 2001.