

FER PREDICTION WITH VARIABLE CODEWORD LENGTH

Alberto Rico-Alvariño, Robert W. Heath Jr.† and Carlos Mosquera**

*Signal Theory & Communications Dept.
University of Vigo, Galicia, Spain
{alberto,mosquera}@gts.uvigo.es

†Dept. of Electrical and Computer Engineering
The University of Texas at Austin, TX, USA
rheath@utexas.edu

ABSTRACT

Frame error rate (FER) prediction in wireless communication systems is an important tool with applications to system level simulations and link adaptation, among others. Although in realistic communication scenarios it is expected to have codewords of different lengths, previous work on FER prediction marginally treated the dependency of the FER on the codeword length. In this paper, we present a method to estimate the FER using codewords of different length. We derive a low complexity FER estimator for frames of different length transmitted over a binary symmetric channel of unknown error probability. We extend this technique to coded systems by the use of effective SNR FER predictors. The proposed estimation scheme is shown to outperform other simpler estimation methods.

Index Terms— FER prediction, PHY abstraction, effective SNR

1. INTRODUCTION

Frame error rate (FER) prediction is an important tool with applications to different communication problems. FER prediction is used for example to provide a PHY layer abstraction for system-level simulations [1–3]. It is also used in link adaptation algorithms where a modulation and coding scheme (MCS) is selected according to the state of the channel [4–7].

FER prediction is challenging in systems employing multiple input multiple output (MIMO) and orthogonal frequency division multiplexing (OFDM) because of the space and frequency selectivity. This selectivity causes the different bits in a codeword to experience different signal to noise ratio (SNR), so the mapping between channel state and FER is more complicated. The problem is compounded by error control coding and interleaving. Different link quality metrics,

link-to-system mappings, or effective SNR metrics [1, 2, 6, 7] have been proposed to perform FER prediction in selective channels. The effective SNR is a Kolmogorov mean [8] of the different SNR values in the different carriers or spatial streams, and usually require the calibration of some parameters according to empirical results [3]. More recently, machine learning methods have been also proposed to perform FER prediction [4, 5, 9, 10].

A limitation of prior work [1, 7] is that the predictions are tuned to frames of a fixed length. Extending these approaches to systems with variable frame length, for example as experienced in wireless local area networks, requires implementing and calibrating for every possible frame length. Unfortunately, this is challenging due to the large number of potential frame sizes. Previous work using FER prediction dealt with the variable codeword length in various ways. For example, [11–14] assumed perfect knowledge of the coded bit error rate (CBER), and from that value they calculated the corresponding FER. Other previous work using FER predictors assumed constant frame length [4–6, 9, 15], which is not realistic under modern communication standards.

In this paper we present a method to perform FER estimation for codewords of variable length. As opposed to [1, 7], we assume the availability of a limited number of FER predictors, calibrated to estimate the FER of a small number of codeword sizes. Unlike [11–14], we account for the FER prediction error of the different codeword sizes. The proposed FER estimator is derived under a binary symmetric channel (BSC) assumption, where the FER estimation is imperfect as a result of having a finite observation window. We extend the result to coded systems under frequency and space selective channels by the use of effective SNR FER predictors. The use of the proposed estimation method enables accurate FER prediction for a wide range of codeword lengths while keeping a low computational complexity.

2. FER ESTIMATION FOR UNCODED SYSTEMS

Consider a transmitter-receiver pair communicating through a noisy channel. The transmitter builds blocks of bits

This work was performed while A. Rico-Alvarino was visiting The University of Texas at Austin. Work by A. Rico-Alvariño and C. Mosquera supported by the Spanish Government under projects DYNACS (TEC2010-21245-C02-02/TCM) and COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010), and the Atlantic Research Center for Information and Communication Technologies (AtlantTIC). Work by A. Rico-Alvariño was supported in part by a Fundacion Pedro Barrie de la Maza graduate scholarship. R. W. Heath Jr. was supported in part by the National Science Foundation under Grant No.1218338 and 1319556.

$[b_1, \dots, b_L]$ of variable length L , with $b_i \in \{0, 1\}$, and the receiver observes $[\hat{b}_1, \dots, \hat{b}_L]$ at the output of the channel, with $\hat{b}_i \in \{0, 1\}$. The channel is memoryless and symmetric with error probability p , i.e., $p \triangleq \mathbb{P}[\hat{b}_i = 1|b_i = 0] = \mathbb{P}[\hat{b}_i = 0|b_i = 1]$. We assume that the transmitted block of bits contains an error detection code such that the receiver is able to identify the received blocks with at least one erroneous bit. We also assume that the error detection code is designed to make the missed detection probability negligible. If we denote by θ_L the probability of receiving an erroneous block of length L , we have that

$$\theta_L = 1 - \mathbb{P}\left[\bigcap_{i=1}^L (b_i = \hat{b}_i)\right] = 1 - (1-p)^L. \quad (1)$$

It can be seen from (1) that if the FER for a frame length L is known perfectly, we can easily obtain the FER for a different frame length \tilde{L} as

$$\theta_{\tilde{L}} = 1 - (1 - \theta_L)^{\tilde{L}/L}. \quad (2)$$

In a realistic scenario, however, it is very unlikely that an exact estimate of the FER is available for any length. In general, the available FER estimate is going to be the result of observing the success and failures of frames of a certain length during a time period. Moreover, observations of frames of different length may improve our FER estimate. In the following, we formalize the problem of estimating the FER from observations of frames of different size.

Consider a communication system with ℓ different frame sizes L_1, \dots, L_ℓ . During an observation period, a receiver observes n_i transmissions of size L_i , out of which m_i are received with errors. m_i is binomially distributed with parameters n_i and θ_{L_i} , i.e.

$$p(m_i; \theta_{L_i}) = \binom{n_i}{m_i} \theta_{L_i}^{m_i} (1 - \theta_{L_i})^{n_i - m_i}. \quad (3)$$

The maximum likelihood estimate (MLE) of θ_{L_i} given the observation m_i is the measured FER, i.e., $\hat{\theta}_{L_i} = \frac{m_i}{n_i}$. The MLE is unbiased with variance

$$\sigma_i^2 \triangleq \mathbb{E}\left[\left(\theta_{L_i} - \hat{\theta}_{L_i}\right)^2\right] = \frac{(1 - \theta_{L_i})\theta_{L_i}}{n_i}. \quad (4)$$

If there are observations from only one length (i.e., $\ell = 1$), we can relate the MLE of θ_{L_1} and $\theta_{\tilde{L}}$ by using the invariance property [16] and (2) as $\hat{\theta}_{\tilde{L}} = 1 - \left(1 - \hat{\theta}_{L_1}\right)^{\tilde{L}/L_1}$.

If $\ell > 1$, however, the derivation of the MLE of $\theta_{\tilde{L}}$ is more involved. For a vector of observed errors $\mathbf{m} \triangleq [m_1, \dots, m_\ell]^T$, the probability mass function of \mathbf{m} parametrized

by the FER $\theta_{\tilde{L}}$ can be easily obtained just by assuming independent observations and by applying (2) as

$$p(\mathbf{m}; \theta_{\tilde{L}}) = \prod_{i=1}^{\ell} \binom{n_i}{m_i} \left(1 - (1 - \theta_{\tilde{L}})^{L_i/\tilde{L}}\right)^{m_i} \times (1 - \theta_{\tilde{L}})^{(n_i - m_i)L_i/\tilde{L}}. \quad (5)$$

It is possible to calculate the Fisher information matrix from (5) and conclude that a minimum variance unbiased estimator of $\theta_{\tilde{L}}$ does not exist [16]. The MLE seems also difficult to calculate, since the likelihood function is nonconcave in $\theta_{\tilde{L}}$ and, therefore, maximizing it would require a grid search. We propose a simple and computationally efficient approach consisting on a linear combination of ℓ MLE of $\theta_{\tilde{L}}$, each one obtained from the observations from a different length. First, let us denote as $\hat{\theta}_{\tilde{L}}(L_i) \triangleq 1 - \left(1 - \hat{\theta}_{L_i}\right)^{\tilde{L}/L_i}$ the MLE of $\theta_{\tilde{L}}$ from the measurements of length L_i . We propose to estimate $\theta_{\tilde{L}}$ by a linear combination of $\hat{\theta}_{\tilde{L}}(L_1), \dots, \hat{\theta}_{\tilde{L}}(L_\ell)$:

$$\hat{\theta}_{\tilde{L}} = \sum_{i=1}^{\ell} \beta_i \hat{\theta}_{\tilde{L}}(L_i) \quad (6)$$

with $\{\beta_i\}$ a set of weights to be designed. Some simple direct values of β_i will serve as our baseline for comparison:

- **Average:** $\beta_i = 1/\ell \forall i$.
- **Closest:** $\beta_i = \begin{cases} 1 & \text{if } i = \arg \min |L_i - \tilde{L}| \\ 0 & \text{otherwise} \end{cases}$.

Obtaining the optimum weights β_i is quite involved as the MLEs $\hat{\theta}_{\tilde{L}}(L_i)$ are biased in general (even when the MLEs $\hat{\theta}_{L_i}$ are not). Also, the variance of $\hat{\theta}_{\tilde{L}}(L_i)$ depends on the parameter to estimate. In the following, we derive a value for the β_i by assuming a sufficiently large number of observations n_i .

If we have a sufficiently large number of observations n_i , the distribution of the MLE $\hat{\theta}_{L_i}$ can be approximated as [16] $\hat{\theta}_{L_i} \sim \mathcal{N}(\theta_{L_i}, \sigma_i^2)$, so we can write

$$\hat{\theta}_{\tilde{L}}(L_i) = 1 - (1 - \theta_{L_i} + w_i)^{\alpha_i} \quad (7)$$

with $\alpha_i = \tilde{L}/L_i$ and $w_i \sim \mathcal{N}(0, \sigma_i^2)$. As we are using a small variance approximation, we approximate $\hat{\theta}_{\tilde{L}}(L_i)$ by the first order Taylor expansion series around $w_i = 0$ as

$$\hat{\theta}_{\tilde{L}}(L_i) \approx \theta_{\tilde{L}} + w_i \alpha_i (1 - \theta_{\tilde{L}})^{\frac{\alpha_i - 1}{\alpha_i}}. \quad (8)$$

From (8) we can see that the MLE $\hat{\theta}_{\tilde{L}}(L_i)$ is unbiased for large n_i , as predicted by the asymptotic properties of any MLE. The variance of the estimator $\hat{\theta}_{\tilde{L}}(L_i)$ is

$$\begin{aligned} \xi_i^2 &\triangleq \mathbb{E}\left[\left(\hat{\theta}_{\tilde{L}}(L_i) - \theta_{\tilde{L}}\right)^2\right] = \sigma_i^2 \alpha_i^2 (1 - \theta_{\tilde{L}})^{2\frac{\alpha_i - 1}{\alpha_i}} \\ &= \frac{\alpha_i^2 \left(1 - (1 - \theta_{\tilde{L}})^{1/\alpha_i}\right) (1 - \theta_{\tilde{L}})^{\frac{2\alpha_i - 1}{\alpha_i}}}{n_i}. \end{aligned} \quad (9)$$

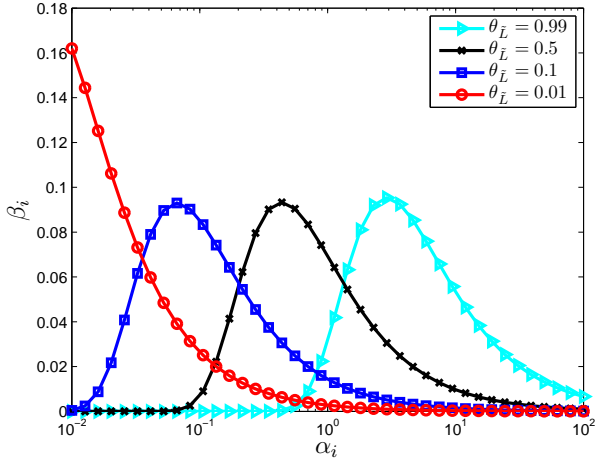


Fig. 1. Linear combination weights β_i as a function of the error probability. We see that, in general, $\alpha_i = 1$ (same length) is not a maximum. Low α values have more weight for low FER values.

From the values ξ_i^2 , the linear fusion weights which minimize the variance of $\hat{\theta}_{\bar{L}}$ in (6) are

$$\beta_i = \frac{1/\xi_i^2}{\sum_{i=1}^{\ell} 1/\xi_i^2}. \quad (10)$$

Note that the MSE ξ_i depends on $\theta_{\bar{L}}$, which is the parameter to be estimated. Thus, the parameters β_i cannot be obtained directly. It is expected, however, that the parameters β_i are not very sensitive to small changes in $\theta_{\bar{L}}$. We propose a two-step estimation. First, obtain an initial estimate $\hat{\theta}_{\bar{L},0}$ of $\theta_{\bar{L}}$ by averaging the MLEs $\hat{\theta}_{\bar{L}}(L_i)$ with weights $\beta_i = 1/\ell$. Second, use $\hat{\theta}_{\bar{L},0}$ as the true value of $\theta_{\bar{L}}$ to calculate the ξ_i and then obtain the optimum weights β_i according to (10).

In Figure 1 we show the evolution of β_i with α_i for different $\theta_{\bar{L}}$, and a constant n_i . It can be seen that the maximum weight is not given to the samples with similar L_i (i.e., $\alpha \approx 1$), and depends on the operating regime. For example, in low FER values ($\theta_{\bar{L}} = 0.01$) more weight is given to samples from longer packets (small α), since in those packets the error probability is going to be larger and, therefore, more error events can be observed. The opposite behavior is observed in the high FER region ($\theta_{\bar{L}} = 0.99$), where shorter packets provide better error estimates.

We evaluated the performance of the proposed estimation approach in a BSC with error probability p . In this setting, the FER is exactly $\theta_L = 1 - (1 - p)^L$, so we can compare the obtained result with the exact FER value. We set $\ell = 5$, with $\mathbf{L} = [L_1, \dots, L_\ell] = [100, 1000, 5000, 8000, 10000]$ and $n_i = N \forall i$, with $N = 10, 100, 1000$. Our objective is to estimate the FER with $\bar{L} = 2000$. We compare the results with the two simpler estimation approaches already

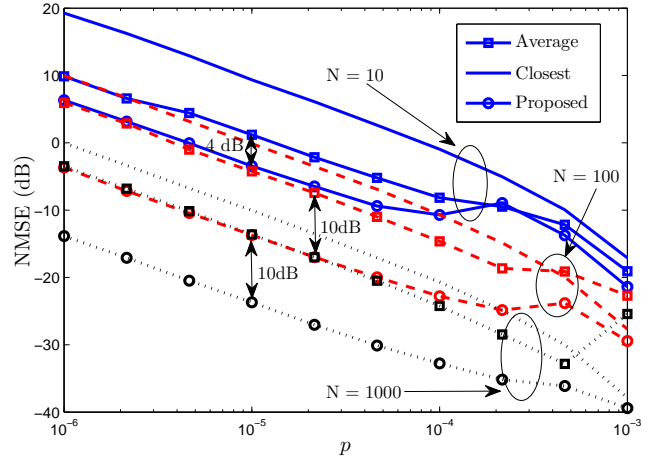


Fig. 2. NMSE of FER prediction with different estimation methods as a function of p . Different colors are used for different codeword lengths: $N = 10$ (blue), $N = 100$ (red), $N = 1000$ (black).

mentioned. The results are shown in Figure 2, where the figure of merit is the normalized mean squared error (NMSE), defined as $\text{NMSE}(dB) = 10 \log_{10} \left(\frac{1}{K} \sum_{i=1}^K \frac{(\hat{\theta}_{\bar{L},i} - \theta_{\bar{L}})^2}{\theta_{\bar{L},i}^2} \right)$ with $\hat{\theta}_{\bar{L},i}$ the FER estimate from the i -th realization of the observations. We averaged the results over $K = 10^4$ realizations of the observations. We can see that the proposed approach outperforms the more naive estimates for almost all values of p and N . The gain is especially noticeable when the number of observations is high ($N = 100, 1000$), and can be as large as 10dB. From the figure, we can observe that the effect in NMSE reduction when applying the proposed method is approximately the same as multiplying the number of observations by a factor of 10, especially for low p values.

We also compared the proposed method against a scenario where all the observations are of the desired length. For example, for $N = 10$, our method observes 10 frames of each length in $\mathbf{L} = [100, 1000, 5000, 8000, 10000]$, and we compare it against the case of observing 50 frames of length 2000. In the latter case, the MLE of the FER is simply the observed FER, and its variance is given by (4). We show the results in Figure 3. The proposed method with observations of different lengths outperforms the MLE with observations of frames of the desired length, especially for low p and large N values.

3. FER ESTIMATION FOR CODED SYSTEMS

We exploit the insights obtained for the BSC to perform FER prediction with different codeword length in coded systems under frequency and space selective channels. Although in general it is difficult to obtain a good approxi-

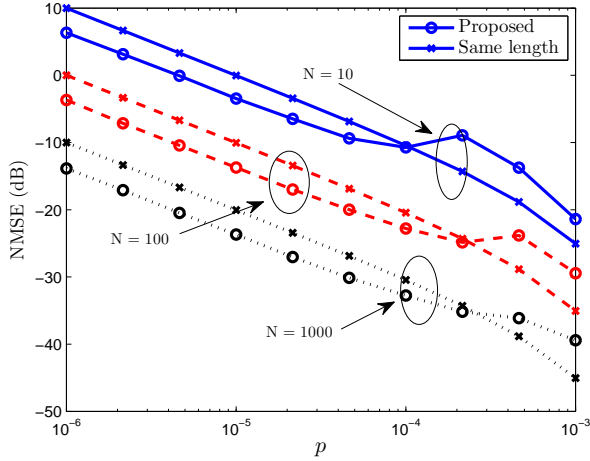


Fig. 3. NMSE of FER prediction compared with observations of the desired length, for different p and N values.

mation for the FER in coded systems, the relationship between the FER and CBER p_{coded} can be approximated as $\theta_L = 1 - (1 - p_{\text{coded}})^{L/D}$, where D is a parameter that depends on the code. For example, D is the minimum distance of the code in the case of convolutional codes, and the number of coded bits in the case of LDPC codes [14]. The relationship between FER of two different lengths is given by (2), thus not depending on the actual value of D .

We perform FER prediction by the use of the exponential effective SNR metric (EESM). Prediction is performed in two steps. First, the set of post-processing SNR values $\gamma_1, \dots, \gamma_K$, are mapped to a single SNR value as

$$\gamma_{\text{eff}} = -\frac{1}{\beta} \log \left(\frac{1}{K} \sum_{i=1}^K e^{-\beta \gamma_k} \right) \quad (11)$$

where β is a calibration parameter to be fitted according to empirical results. Second, the FER estimate for a certain length L is obtained as $\hat{\theta}_L = \text{FER}_{\text{AWGN},L}(\gamma_{\text{eff}})$ with $\text{FER}_{\text{AWGN},L}$ an empirically obtained function mapping an SNR value to the FER in AWGN for a codeword length L . Note that both the calibration parameter β and the FER in AWGN are going to depend on L . Also, if FER prediction is performed for different modulation and coding schemes, a different β and FER function has to be obtained for each case.

We evaluated the proposed estimation procedure for FER prediction under the IEEE 802.11ac standard [17]. We selected MCS QPSK 3/4 with two spatial streams, a transmitter-receiver pair with 2 antennas each, and SVD precoding. We obtained FER samples for codewords with length $\mathbf{L} = [2^7, \dots, 2^{14}]$, i.e., codeword lengths between 128 and 16384 bits. We trained 8 EESM estimators (one for each length) with 500 FER samples for different SNR values and

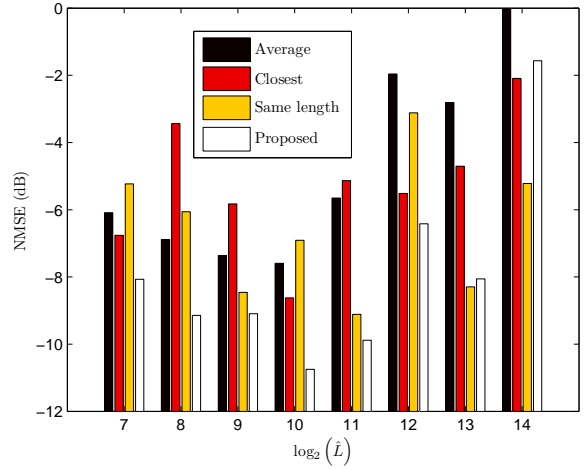


Fig. 4. NMSE of FER prediction for different frame sizes in IEEE 802.11ac, QPSK 3/4.

Rayleigh channel model. The SNR values were selected between 0 and 30 dB, so FER values between 0 and 1 had to be estimated. For each of the 8 length values, we estimated the FER using the other 7 FER predictors (with the 3 proposed linear fusion rules), and compared the result with the FER prediction from the EESM estimator of codewords of the same length. Formally, the proposed FER estimator is

$$\hat{\theta}_{\tilde{L}} = \sum_{i=1, L_i \neq \tilde{L}}^{\ell} \beta_i (1 - (1 - \text{FER}_{\text{AWGN},L_i}(\gamma_{\text{eff}}))^{\alpha_i}). \quad (12)$$

We obtained the designed weights β_i by assuming a constant n_i in (9). We show the results in Figure 4. We can see that the proposed estimator outperforms the other linear fusion approaches for all cases, except for the higher length case, where it attains approximately the same NMSE as the *Closest* combining. Also, for 6 of the 8 lengths, the proposed estimator outperforms the EESM FER prediction trained with samples of the same length. In some cases, like $\tilde{L} = 2^{10}$, the NMSE gain with respect to the same length FER predictor is in the order of 4dB.

4. CONCLUSIONS

In this paper we presented an approach to perform FER estimation in communication systems using codewords of different length. We derived an estimator as a linear combination of several MLE of the FER in a BSC. We use this estimator in a coded system with FER prediction using EESM, and we compare its performance against other FER prediction approaches. The proposed estimator with the designed linear combination is shown to outperform other simpler estimation methods. We also show that a MLE with samples of the desired length can perform worse than the proposed estimator.

5. REFERENCES

- [1] Y. Blankenship, P. Sartori, B. Classon, V. Desai, and K. Baum, "Link error prediction methods for multicarrier systems," in *Proc. IEEE VTC Fall*, vol. 6, (Los Angeles, CA), pp. 4175–4179 Vol. 6, Sept. 2004.
- [2] K. Brueninghaus, D. Astely, T. Salzer, S. Visuri, A. Alexiou, S. Karger, and G.-A. Seraji, "Link performance models for system level simulations of broadband radio access systems," in *Proc. IEEE PIMRC*, vol. 4, (Berlin, Germany), pp. 2306–2311 Vol. 4, Sept. 2005.
- [3] A. Cipriano, R. Visoz, and T. Salzer, "Calibration issues of PHY layer abstractions for wireless broadband systems," in *Proc. IEEE VTC Fall*, (Calgary, Canada), pp. 1–5, 2008.
- [4] S. Wahls and H. V. Poor, "Link adaptation for BICM-OFDM through adaptive kernel regression," in *Proc. IEEE ICASSP 2013*, (Vancouver, Canada), pp. 5136–5140, 2013.
- [5] Z. Puljiz, M. Park, and R. W. Heath Jr., "A machine learning approach to link adaptation for SC-FDE system," in *Proc. IEEE GLOBECOM*, (Houston, TX), pp. 1–5, Dec. 2011.
- [6] M. Lampe, T. Giesel, H. Rohling, and W. Zirwas, "PER prediction for PHY mode selection in OFDM communication systems," in *Proc. IEEE GLOBECOM*, vol. 1, (San Francisco, CA), pp. 25–29 Vol.1, 2003.
- [7] T. Jensen, S. Kant, J. Wehinger, and B. Fleury, "Mutual information metrics for fast link adaptation in IEEE 802.11n," in *Proc. IEEE ICC*, (Beijing, China), pp. 4910–4915, May 2008.
- [8] A. Kolmogorov, *Sur la notion de la moyenne*. Atti della R. Accad. dei Lincei, 1930.
- [9] R. Daniels, C. Caramanis, and R. W. Heath Jr., "Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and SNR ordering," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 114–126, Jan. 2010.
- [10] A. Rico-Alvarino and R. W. Heath Jr, "Learning-based adaptive transmission for limited feedback multiuser MIMO-OFDM," *Submitted to IEEE Transactions on Wireless Communications*, June 2013. Available at <http://arxiv.org/abs/1306.0502>.
- [11] C. Wang, "Dynamic ARF for throughput improvement in 802.11 WLAN via a machine-learning approach," *Journal of Network and Computer Applications*, vol. 36, no. 2, pp. 667 – 676, 2013.
- [12] G. Martorell, F. Riera-Palou, and G. Femenias, "Cross-layer fast link adaptation for MIMO-OFDM based WLANs," *Wireless Personal Communications*, vol. 56, no. 3, pp. 599–609, 2011.
- [13] D. Qiao, S. Choi, and K. Shin, "Goodput analysis and link adaptation for IEEE 802.11a wireless LANs," *IEEE Trans. on Mobile Computing*, vol. 1, no. 4, pp. 278–292, 2002.
- [14] P. Tan, Y. Wu, and S. Sun, "Link adaptation based on adaptive modulation and coding for multiple-antenna OFDM system," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1599–1606, 2008.
- [15] Q. Liu, S. Zhou, and G. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1746 – 1755, Sept. 2004.
- [16] S. Kay, "Fundamentals of statistical signal processing, volume 1," *Signal Processing*, 1993.
- [17] "IEEE draft standard for IT - telecommunications and information exchange between systems - LAN/MAN - specific requirements - part 11: Wireless LAN medium access control and physical layer specifications - amd 4: Enhancements for very high throughput for operation in bands below 6GHz," *IEEE P802.11ac/D3.0*, June 2012, pp. 1–385, jun 2012.