Adaptive Modulation Systems for Predicted Wireless Channels

Sorour Falahati[1], Arne Svensson[1], Torbjörn Ekman[2] and Mikael Sternad[3]

[1]Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Göteborg, Sweden.

Email: {Sorour.Falahati, Arne.Svensson}@s2.chalmers.se

[2]Unik, P.O.Box 70, N-2027, Kjeller, Norway.

Email: torbjorn@unik.no

[3]Signals and Systems, Department of Material Science, Uppsala University, SE-751 20 Uppsala, Sweden.

Email: Mikael.Sternad@signal.uu.se

Abstract

When adaptive modulation is used to counter short-term fading in mobile radio channels, signaling delays create problems with outdated channel state information. The use of channel power prediction will improve the performance of the link adaptation. It is then of interest to take the quality of these predictions into account explicitly when designing an adaptive modulation scheme. We study the optimum design of an adaptive modulation scheme based on uncoded M-QAM modulation assisted by channel prediction for the flat Rayleigh fading channel. The data rate, and in some variants the transmit power, are adapted to maximize the spectral efficiency subject to average power and bit error rate constraints. The key issues studied here are how a known prediction error variance will affect the optimized transmission properties such as the SNR boundaries that determine when to apply different modulation rates, and to what extent it affects the spectral efficiency. This investigation is performed by analytical optimization of the link adaptation, using the statistical properties of a particular but efficient channel power predictor. Optimum solutions for the rate and transmit power are derived based on the predicted SNR and the prediction error variance. The numerical results are complemented by simulations.

1 Introduction

Considering the rapidly growing demand for mobile communications while limited spectrum is available, spectrally efficient communication techniques are of great importance in future wireless communications. Adaptive modulation, or link adaptation, is a powerful technique for improving the spectral efficiency in wireless transmission over fading channels. If complete Channel State Information (CSI) is known at the transmitter, the Shannon capacity of a fading channel can be approached by optimal adaptation of the signaling parameters such as transmit power, data rate, channel coding rate or scheme [1,2]. Adaptive modulation has been extensively studied in [3–14] and the references therein.

With the adaptive modulation considered here, a high spectral efficiency is achievable at a given Bit Error Rate (BER) in favorable channel conditions, while a reduction of the throughput is experienced when the channel degrades. The adaptation can also take requirements of different traffic classes and services such as required BERs, into account.

We consider *fast* link adaptation, i.e. we strive to adapt to the small scale fading. The receiver estimates the received power and sends feedback information via a return channel to the transmitter, with the aim of modifying the modulation parameters. Due to the unavoidable delays involved in power estimation, feedback transmission and modulation adjustment, the fast link adaptation needs to be based on *predicted* estimates of the power of the fading communication channel. In the so far proposed solutions for optimum design of adaptive modulation systems, perfect knowledge of the CSI at the transmitter as well as error free channel estimates at the receiver are common assumptions for the system design and performance evaluation. In real systems, these assumptions are not valid. Due to the time-varying nature of the wireless channels, the channel status will change during the time delay between estimation and data transmission. This leads to performance degradation such as decrease in the throughput.

The impact of the uncertainty in channel estimates on the performance has been discussed in the literature (see e.g. [5–7,15–20]). In [17], the effect of channel estimation errors on the BER performance is illustrated. However, this effect is investigated on the receiver side, not on the choice of link adaptation parameters. In [5,6], the impact of time delay on the adaptive modulation performance is characterized. It is shown that systems with low BER requirements are very sensitive to the time delay. They can operate acceptably only if the delay is kept below a critical small value. In [7], a linear predictor is used to estimate the current channel status based on the outdated estimates. The channel status is used to determine the currently appropriate modulation, and the effect of the time delay and mobile speed on the BER and throughput performance are studied. However, the SNR thresholds which determine the modulation modes are evaluated based on the simulation results only.

Results by Goeckel [15,16] highlight that to design a practical adaptive modulation system, time variations of the channel which severely degrade the system performance, should be taken into account. There, a novel approach to the design of a robust adaptive modulation system based on only a single outdated fading estimate is proposed. In [18–20], adaptive modulation schemes based on accurate long-rang CSI prediction are investigated.

The system proposed here utilizes an unbiased quadratic regression of historic noisy channel estimates to predict the signal power at the receiver [21,22]. For this type of predictor, there exists a statistical model for the prediction error for Rayleigh fading channels which enables an analytical optimization of the rate adaptation scheme. This statistical model will also be used for analyzing the resulting BER and spectral efficiency for given prediction error variances. We restrict our attention to link adaptation with uncoded M-QAM modulation. With no coding, the two remaining degrees of freedom are the choice of modulation formats in different SNR regions, and the possibility to use transmit power control within these regions. Exploitation of the statistical information about the prediction errors will be shown to improve the overall system

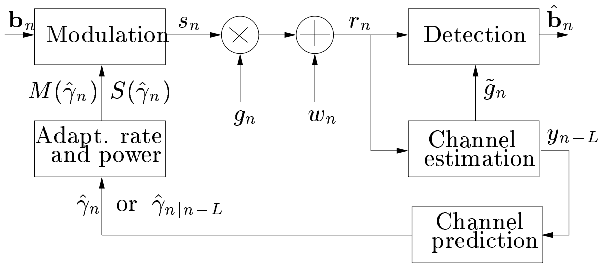

Figure 1: Discrete model of the system.

performance, and adjusts the link adaptation better to the true channel conditions. As a result, the BER constraints will be fulfilled in the presence of prediction errors. This will not be the case if the prediction errors are neglected in the design of the link adaptation.

This paper is organized as follows. Section 2 describes the system model and the notations which are used throughout this study. The channel prediction is explained in Section 3. The BER is evaluated as a function of predicted instantaneous SNR in Section 4 and optimal rate and power adaptation are derived under different constraints in Section 5. Analytical results are presented in Section 6 while Section 7 summarizes the results.

2 System Model

In the adaptive modulation scheme, M-QAM modulation schemes with different constellation sizes are provided at the transmitter. For each transmission, the modulation scheme and possibly also the transmit power are adjusted to maximize the spectral efficiency, under BER and average power constraints, based on the instantaneous predicted SNR. The channel is modelled by a flat Rayleigh fading channel. At the receiver, demodulation is performed using channel estimates. The discrete model of the system is depicted in Figure 1. All the signals are sampled at the symbol rate where the index n represents the signal sample at time nT_s where T_s is the symbol period. Here, g_n is the zero mean, complex channel gain with circular Gaussian distribution where the power $|g_n|^2$ is χ^2 distributed. The auto-correlation function of the complex channel gain is assumed given by Jakes model

$$r_g(m) = E(g_n g_{n-m}^*) = \sigma_g^2 J_0(2\pi f_D T_s m) \qquad (1)$$

where σ_g^2 is the average power of the complex channel gain, $E(\ )$ is the expectation function, the asterisk denotes the conjugate operator, $J_0(\ )$ is the Bessel function of order zero and first kind, and f_D is the maximum Doppler frequency. In the following, r_g will frequently be used to denote $r_g(0) = \sigma_g^2$. Moreover, w_n is a sample of a complex Additive White Gaussian Noise (AWGN) with zero mean and time-invariant variance σ_w^2. The estimate y_{n-L} is the noisy observation of g_{n-L} at the receiver. A time-series of these estimates are used at the receiver to predict the channel power gain $|g_n|^2$ which is proportional to the instantaneous received SNR, denoted by γ_n. Either the predicted SNR, denoted by $\hat{\gamma}_{n|n-L}$ or $\hat{\gamma}_n$, or the corresponding appropriate rate and transmission power levels are then fed back to the transmitter[1]. An error free feed-back channel is assumed. The prediction horizon L is assumed to be larger than the sum of computational delays and signalling delays in the adaptation control loop. The SNR prediction $\hat{\gamma}_n$ is here assumed to be based on noisy channel estimates y_{n-L} and will therefore have a prediction error.

Based on $\hat{\gamma}_n$, a modulation scheme with constellation size $M(\hat{\gamma}_n)$ (out of N constellations available at the transmitter), which transmits $k(\hat{\gamma}_n) = \log_2 M(\hat{\gamma}_n)$ bits per symbol, and a transmit power $S(\hat{\gamma}_n)$ are selected. Each block of $k(\hat{\gamma}_n)$ data bits denoted by $\mathbf{b}_n$, is Gray encoded and mapped to a symbol in the signal constellation denoted by s_n, which is transmitted over the flat Rayleigh fading channel. The received sample, r_n, is used to estimate the channel gain $\tilde{g}_n$, which in turn is used to detect the transmitted bits denoted by $\hat{\mathbf{b}}_n$. Since the estimation error is believed to have a minor effect on the performance compared to the prediction error, perfect channel estimation is here assumed for the demodulation, i.e. $\tilde{g}_n = g_n$.

In this study, the following notations similar to those of [14] are used. Let $\bar{S}$ denote the average transmit signal power. The average received SNR is then given by

$$\bar{\gamma} = r_g \frac{\bar{S}}{\sigma_w^2}. \tag{2}$$

For a constant transmit power $\bar{S}$, the instantaneous received SNR is

$$\gamma_n = \bar{\gamma} \frac{p_n}{r_g} \tag{3}$$

[1] In practice, the feedback information will have to be quantized to limit the return channel bandwidth. This added source of error is not taken into account in the analysis.

where $p_n = |g_n|^2$ is the instantaneous channel power gain. Accordingly, the instantaneous predicted received SNR is

$$\hat{\gamma}_n = \bar{\gamma} \frac{\hat{p}_{n|n-L}}{r_g} \tag{4}$$

where $\hat{p}_{n|n-L}$ is the predicted instantaneous channel power gain $|g_n|^2$. For the transmit power $S(\hat{\gamma}_n)$, the instantaneous received SNR is given by $\gamma_n(S(\hat{\gamma}_n)/\bar{S})$, with γ_n given by (3).

The rate region boundaries, defined as the ranges of $\hat{\gamma}_n$ values over which the different constellations are used by the transmitter, are denoted by $\{\hat{\gamma}_i\}_{i=0}^{N-1}$. When the predicted instantaneous SNR belongs to a given rate region, i.e. $\hat{\gamma}_n \in [\hat{\gamma}_i, \hat{\gamma}_{i+1})$, the corresponding constellation of size $M(\hat{\gamma}_n) = M_i$ with $k(\hat{\gamma}_n) = k_i$ bits per symbol is transmitted where $\hat{\gamma}_N = \infty$. There is no transmission if $\hat{\gamma}_n < \hat{\gamma}_0$, meaning that $\hat{\gamma}_0$ is the cutoff SNR.

3 Channel Prediction

The channel gain in a Rayleigh fading mobile radio channel takes values from a complex time series which can be modelled as a correlated complex Gaussian stochastic process. The absolute square, i.e. the power, of the time series is here assumed to be predicted based on previous observations of the complex time series y_n. The observed complex channel gain is assumed to be affected by an additive estimation error e_n where e_n is assumed to be a white and zero mean complex Gaussian random variable which is independent of g_n.[2] Thus, $y_n = g_n + e_n$.

Based on a finite number of past observations of y_n, the complex channel at time n could be predicted with a prediction horizon L by a linear FIR filter

$$\hat{g}_{n|n-L} = \varphi_{n-L}^H \theta \tag{5}$$

where θ is a column vector containing K complex-valued predictor coefficients and

$$\varphi_{n-L}^H = [y_{n-L}, y_{n-L-1}, \cdots, y_{n-L-(K-1)}], \tag{6}$$

[2] Theoretically, Rayleigh fading channels with autocorrelation function given by (1), are band-limited stochastic processes, and should therefore be perfectly predictable for arbitrarily long prediction horizons. However, since only noisy observations of the channel are available, predictability will decrease with an increasing horizon. The predictor performance will depend on the use of good channel estimates and the application of noise-reducing smoothing, which improve $E|g_n|^2/E|e_n|^2$, see [21–23].

is the regressor where H represents a Hermitian transpose. A Wiener adjustment of θ provides the optimal linear predictor in the Mean Square Error (MSE) sense.

The adjustment of an adaptive modulation scheme is determined not by the complex channel gain g_n, but by the SNR at the time of transmission. If we, for simplicity, assume that the variance of the noise and disturbance w_n in Figure 1 is constant, the channel power $p_n = |g_n|^2$ will have to be predicted. However, the use of the squared magnitude of the linear prediction $\hat{g}_{n|n-L}$ as a predictor of the channel power would on average underestimate the true power, and result in a biased estimate. The reason is that the average power of $\hat{g}_{n|n-L}$ will decrease with an increasing prediction horizon L and be lower than the average power of g_{n+L}, due to the limited predictability of the process g_n. We here instead utilize a recently developed quadratic power predictor which eliminates this bias. It is given by

$$\hat{p}_{n|n-L} = \theta^H \varphi_{n-L} \varphi_{n-L}^H \theta + r_g - \theta^H \mathbf{R}_\varphi \theta. \qquad (7)$$

Here, $\mathbf{R}_\varphi = E(\varphi_{n-L}\varphi_{n-L}^H)$ is the $K \times K$ correlation matrix for the regressors. Note that $E(\hat{p}_{n|n-L}) = r_g$ for all L. The unbiased quadratic predictor that minimizes the power MSE is derived in [21], where it is shown that the predictor coefficient vector θ that provides an MSE optimal the channel predictor (5) will *also* result in an MSE optimal power predictor, when used in (7). The optimal adjustment for both of these problems is thus given by

$$\theta = \mathbf{R}_\varphi^{-1} \mathbf{r}_{g\varphi} \qquad (8)$$

where

$$\begin{aligned} \mathbf{r}_{g\varphi} &= \mathrm{E}\{g_n \varphi_{n-L}\} \\ &= [r_g(L), r_g(L+1), \cdots, r_g(L+(K-1))]^T. \end{aligned} \qquad (9)$$

The miminum MSEs of the channel gain prediction error (i.e. $\epsilon_{c_n} = g_n - \hat{g}_{n|n-L}$) and the power prediction error (i.e. $\epsilon_{p_n} = p_n - \hat{p}_{n|n-L}$) are given by

$$\sigma_{\epsilon_c}^2 = r_g - \mathbf{r}_{g\varphi}^H \mathbf{R}_\varphi^{-1} \mathbf{r}_{g\varphi}, \qquad (10)$$

$$\sigma_{\epsilon_p}^2 = r_g^2 - |\mathbf{r}_{g\varphi}^H \mathbf{R}_\varphi^{-1} \mathbf{r}_{g\varphi}|^2, \qquad (11)$$

respectively. Thus, by (8) and (7), the optimum quadratic power predictor can be expressed in terms of the MSE-optimal linear FIR channel predictor as

$$\hat{p}_{n|n-L} = |\hat{g}_{n|n-L}|^2 + \sigma_{\varepsilon_c}^2. \qquad (12)$$

In other words, the squared magnitude of the optimal FIR power prediction is modified simply by adding the variance of this estimate. The bias compensation will reduce the total prediction MSE. It also provides superior performance compared to the use of linear power predictors that are based on channel power samples ($|y_n|^2$) as regressors [22].[3] The assumption behind (10) and (11) is that the unbiased quadratic power predictor is *optimal* in the sense that the second order statistics of g_n have been estimated perfectly, so that the parameter vector θ is perfectly adjusted. Then, given a prediction $\hat{p}_{n|n-L}$, the conditional power prediction error variance, denoted by $\sigma_{\epsilon_{p_c}}^2(\hat{p}_{n|n-L})$, is given by [22]

$$\sigma_{\epsilon_{p_c}}^2(\hat{p}_{n|n-L}) = \sigma_{\epsilon_c}^2 [2\hat{p}_{n|n-L} - \sigma_{\epsilon_c}^2]. \qquad (13)$$

If we average over the predicted power in (13), we obtain

$$\sigma_{\epsilon_p}^2 = \sigma_{\epsilon_c}^2 [2r_g - \sigma_{\epsilon_c}^2], \qquad (14)$$

since, with the unbiased predictor, $E(\hat{p}_{n|n-L}) = E(p_n) = r_g$.

Another indication of the predictor performance is the relative standard deviation of the conditional power prediction error. Using (13), this measure is given by

$$\frac{\sigma_{\epsilon_{p_c}}(\hat{p}_{n|n-L})}{\hat{p}_{n|n-L}} = \sigma_{\epsilon_c} \sqrt{\frac{2\hat{p}_{n|n-L} - \sigma_{\epsilon_c}^2}{\hat{p}_{n|n-L}^2}} \qquad (15)$$

which increases when $\hat{p}_{n|n-L}$ becomes small, i.e. when we predict into a fading dip.

In order to solve the rate adaptation optimization problem, the pdf of the instantaneous SNR is required. In section 8 of [22], it is shown that if an optimal unbiased power predictor is used which provides a given $\sigma_{\epsilon_c}^2 / r_g$, the pdf of γ_n conditioned on $\hat{\gamma}_{n|n-L}$ will be given by

$$\begin{aligned} f_\gamma(\gamma|\hat{\gamma}) = & \frac{\mathrm{U}(\gamma)\mathrm{U}(\hat{\gamma} - \bar{\gamma}\sigma_{\epsilon_c}^2/r_g)}{\bar{\gamma}\sigma_{\epsilon_c}^2/r_g} \times \\ & \exp\left[-\frac{\gamma + \hat{\gamma} - \bar{\gamma}\sigma_{\epsilon_c}^2/r_g}{\bar{\gamma}\sigma_{\epsilon_c}^2/r_g}\right] \times \\ & \mathrm{I}_0\left(\frac{2}{\bar{\gamma}\sigma_{\epsilon_c}^2/r_g}\sqrt{\gamma(\hat{\gamma} - \bar{\gamma}\sigma_{\epsilon_c}^2/r_g)}\right), \end{aligned} \qquad (16)$$

where U( ) is the Heaviside's step function, $\bar{\gamma}$ is given by (2) and I_0( ) is the zeroth order modified

[3] To provide the best prediction performance on fading channels for a limited pre-specified number K of parameters, the delay spacing (the time delay between consecutive channel samples in the regressor) should be selected in a way appropriate for the fading rate.

Bessel function. The time index n was dropped in the pdf expressions since γ_n and $\hat{\gamma}_{n|n-L}$ are both stationary random processes. The pdf of $\hat{\gamma}$ will be given by

$$f_{\hat{\gamma}}(\hat{\gamma}) = \frac{\mathrm{U}(\hat{\gamma} - \bar{\gamma}\sigma^2_{\epsilon_c}/r_g)}{\bar{\gamma}(1-\sigma^2_{\epsilon_c}/r_g)} \exp\left[-\frac{\hat{\gamma} - \bar{\gamma}\sigma^2_{\epsilon_c}/r_g}{\bar{\gamma}(1-\sigma^2_{\epsilon_c}/r_g)}\right]. \tag{17}$$

This is a shifted $\chi^2(2)$-distribution, with the shift $\bar{\gamma}\sigma^2_{\varepsilon_c}/r_g$ caused by the bias compensation term in (7) and (12).

4 M-QAM BER Performance

The transmitter adjusts the constellation size and possibly also the transmit power based on the instantaneous predicted SNR $\hat{\gamma}_{n|n-L}$, where the time index n will be dropped in the following. Evaluation of the optimal power and constellation size (or rate) adjustments which maximize the spectral efficieny and satisfy the BER requirement, requires an invertible expression for the BER as a function of $\hat{\gamma}$. Assuming a square M-QAM with Gray encoded bits, constellation size M_i, and transmit power $S(\hat{\gamma})$, the instantaneous BER as a function of γ and $\hat{\gamma}$ on an AWGN channel, is approximated by [24]

$$\mathrm{BER}(\gamma, \hat{\gamma}) \approx \frac{2\left(1 - \frac{1}{\sqrt{M_i}}\right)}{\log_2 M_i} \mathrm{erfc}\left(\sqrt{\frac{1.5\gamma \frac{S(\hat{\gamma})}{\bar{S}}}{M_i - 1}}\right) \tag{18}$$

which is tight for high SNRs. In [14], it is shown that (18) can be further approximated as

$$\mathrm{BER}(\gamma, \hat{\gamma}) \approx 0.2 \exp\left(\frac{-1.6\gamma}{M_i - 1} \frac{S(\hat{\gamma})}{\bar{S}}\right) \tag{19}$$

which is tight within 1 dB for $M_i \geq 4$ and BER$\leq 10^{-3}$. By averaging (19) over the whole range of the instantaneous true SNR, γ,

$$\mathrm{BER}(\hat{\gamma}) = \int_0^\infty \mathrm{BER}(\gamma, \hat{\gamma}) f_\gamma(\gamma|\hat{\gamma}) d\gamma, \tag{20}$$

the instantaneous BER as a function of the instantaneous predicted SNR[4], $\hat{\gamma}$, is obtained as

$$\mathrm{BER}(\hat{\gamma}) \approx 0.2 z(\hat{\gamma}) \exp\left[(1 - x(\hat{\gamma}))(1 - z(\hat{\gamma}))\right] \tag{21}$$

[4]This is an average over the pdf over the true instantaneous SNR for one specific modulation scheme, and that the term *instantaneous* refers to the fact that the BER is a function of instantaneous predicted SNR.

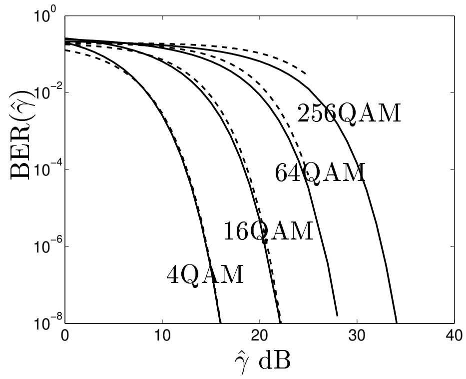

Figure 2: BER versus instantaneous predicted SNR of M-QAM schemes for $\bar{\gamma} = 30$ dB and $\sigma^2_{\epsilon_p} = 0.001$. The solid lines and dashed lines correspond to (20) based on (18) and (19), respectively.

where

$$x(\hat{\gamma}) = \frac{\hat{\gamma}}{\bar{\gamma}\sigma^2_{\epsilon_c}/r_g} \tag{22}$$

and

$$z(\hat{\gamma}) = \frac{1}{1 + A_i S(\hat{\gamma})} \tag{23}$$

where

$$A_i = \frac{1.6}{M_i - 1} \frac{\bar{\gamma}\sigma^2_{\epsilon_c}/r_g}{\bar{S}}. \tag{24}$$

Note that $x(\hat{\gamma}) \geq 1$ and $0 \leq z(\hat{\gamma}) \leq 1$. In Figure 2, the instantaneous BER is illustrated for the constant transmit power $S(\hat{\gamma}) = \bar{S}$, $\bar{\gamma} = 30$ dB and different constellation sizes where both (18) and (19) are used in (20). It is shown that the results (21) based on the approximation given by (19) are tight enough. We have observed the same results for other values of $\bar{\gamma}$ which were examined. Finally, similar to [14], the average BER is given by

$$\overline{\mathrm{BER}} = \frac{\sum_{i=0}^{N-1} k_i \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} \mathrm{BER}(\hat{\gamma}) f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma}}{\sum_{i=0}^{N-1} k_i \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma}}. \tag{25}$$

5 Optimal Rate and Power Adaptation

The spectral efficiency of a modulation scheme is given by the average data rate per unit bandwidth (R/B) where R is the data rate and B is the transmitted signal bandwidth. When a modulation with

constellation size M_i is chosen, the instanteneous data rate is k_i/T_s (bps).
Assuming the Nyquist data pulses ($B = 1/T_s$), the spectral efficiency is given by

$$\eta_B = \frac{R}{B} = \sum_{i=0}^{N-1} k_i \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma} \quad \text{bps/Hz}. \tag{26}$$

As explained in Section 2, a signal constellation and consequently a transmission rate is assigned to each rate region boundary. The rate region boundaries will be adjusted to optimize the spectral efficiency, subject to various constraints. For SNR levels between the rate boundaries, the transmit power may furthermore be adjusted. We thus have an optimization problem with two possible degrees of freedom: the rate region boundaries and the transmit power.

In this work, we consider the following scenarios. First, we intend to maximize the spectal efficiency where both the average power and instantaneous BER are constrained. The transmit power as well as the rate are adapted to satisfy the requirements. Then, we study this case where constant transmit power is presumed. The motivation is that the data rate adaptation has the major effect in increasing the spectral efficiency as compared to the power adaptation, as shown by [5,12]. Also, transmission with variable power complicates the practical implementation: transmission of only integers k_i requires less feedback bandwidth compared to using fast adaptive modulation combined with fast power control. We thereafter relax the BER constraint by constrainting the average BER instead of the instantaneous BER, which results in an increase in the spectral efficiency.

We illustrate the derivation of the optimum rate region boundaries and possibly also transmit power adjustment of these cases. Once the optimal rate region boundaries are evaluated, the spectral efficiency can easily be obtained according to (26).

5.1 Instantaneous BER and variable power (I-BER, V-Pow)

The case we consider first is maximizing the spectral efficiency subject to the average transmit power constraint

$$\int_0^\infty S(\hat{\gamma}) f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma} \leq \bar{S} \tag{27}$$

and the instantaneous BER constraint

$$\text{BER}(\hat{\gamma}) = P_b. \tag{28}$$

The constraint (28) together with (21) show that one of the variables, i.e. $z(\hat{\gamma})$ or $x(\hat{\gamma})$, can be expressed in terms of the other one. Hence, we take the natural logarithm of (28) based on (21) and then use the Taylor approximation $\ln z(\hat{\gamma}) \approx z(\hat{\gamma}) - 1$ about $z(\hat{\gamma}) = 1$ to obtain

$$z(\hat{\gamma}) \approx 1 - \frac{1}{x(\hat{\gamma})} \ln(0.2/P_b) \tag{29}$$

where by substituting (22) and (23) in the above equation, we obtain an expression for the power adjustment within the SNR region for rate i given by

$$\begin{aligned} S_i(\hat{\gamma}) \approx & \left[\frac{\frac{1}{A_i} \frac{\bar{\gamma}\sigma_{\epsilon_c}^2}{r_g} \ln(0.2/P_b)}{\hat{\gamma} - \frac{\bar{\gamma}\sigma_{\epsilon_c}^2}{r_g} \ln(0.2/P_b)} \right] \times \\ & \mathrm{U}\left(\hat{\gamma} - \frac{\bar{\gamma}\sigma_{\epsilon_c}^2}{r_g} \ln(0.2/P_b) \right) \end{aligned} \tag{30}$$

where

$$S_i(\hat{\gamma}) = S(\hat{\gamma}), \quad \hat{\gamma} \in [\hat{\gamma}_i, \hat{\gamma}_{i+1}). \tag{31}$$

This simplifies the optimization problem to finding only the optimal rate region boundaries. Hence, we form the Langrangian function from the spectral efficiency criterion (26) and the power constraint (27), which is here treated as an equality constraint. It is given by

$$\begin{aligned} J(\hat{\gamma}_0, \hat{\gamma}_1, \cdots, \hat{\gamma}_{N-1}) = & \sum_{i=0}^{N-1} k_i \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma} \\ & + \lambda \left(\sum_{i=0}^{N-1} \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} S_i(\hat{\gamma}) f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma} - \bar{S} \right) \end{aligned} \tag{32}$$

where $\lambda \neq 0$ is the Lagrangian multiplier. Solving

$$\frac{\partial J}{\partial \hat{\gamma}_i} = 0, \quad 0 \leq i \leq N-1 \tag{33}$$

results in

$$S_{i-1}(\hat{\gamma}_i) - S_i(\hat{\gamma}_i) = \frac{k_i - k_{i-1}}{\lambda} \tag{34}$$

where $k_{-1} = 0$ and $S_{-1}(\hat{\gamma}) = 0$. From (30) and (34), we obtain

$$\begin{aligned} \hat{\gamma}_i = \ln\left(\frac{0.2}{P_b}\right) \left(\frac{\bar{\gamma}\sigma_{\epsilon_c}^2}{r_g} - \frac{\bar{S}}{1.6} \frac{M_i - M_{i-1}}{k_i - k_{i-1}} \lambda \right), \\ 0 \leq i \leq N-1. \end{aligned} \tag{35}$$

The Lagrange multiplier λ is numerically evaluated based on the power constraint (27). Given (30) and

(35), the average power constraint (27) can be written as

$$\sum_{i=0}^{N-1} \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} S_i(\hat{\gamma}) f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma} = \\ \rho \exp\left(\frac{\sigma_{\epsilon_c}^2/r_g}{1-\sigma_{\epsilon_c}^2/r_g}\left(1-\ln(0.2/P_b)\right)\right) \times \\ \sum_{i=0}^{N-2} (M_i - 1)\left(\mathrm{Ei}(\rho\lambda P_i) - \mathrm{Ei}(\rho\lambda P_{i+1})\right) + \\ (M_{N-1} - 1)\mathrm{Ei}(\rho\lambda P_{N-1}) \leq \bar{S} \tag{36}$$

where Ei( ) is the Exponential Integral and for convenience, the notations

$$\rho = -\ln\left(\frac{0.2}{P_b}\right)\frac{\bar{S}}{1.6}\frac{1}{\bar{\gamma}(1-\sigma_{\epsilon_c}^2/r_g)}, \\ P_i = \frac{M_{i-1} - M_i}{k_{i-1} - k_i} \tag{37}$$

are used. Here, the average power constraint will be fulfilled if $\lambda < 0$. A *bisection method* is used to numerically search for λ which meets the power constraint with equality.

5.2 Instantaneous BER and constant power (I-BER, C-Pow)

The instantaneous BER constraint with constant transmit power is considered here. Due to the constant transmit power i.e. $S(\hat{\gamma}) = S$, we have

$$\mathrm{BER}(\hat{\gamma}) = \frac{0.2}{1+A_iS}\exp\left[\frac{A_iS}{1+A_iS}\left(1-x(\hat{\gamma})\right)\right]. \tag{38}$$

The average power constraint (27), implies that the cut off SNR $\hat{\gamma}_0$, should satisfy

$$\frac{S}{\bar{S}} = \frac{1}{\int_{\hat{\gamma}_0}^{\infty} f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma}} \tag{39}$$

which leads to

$$S = \bar{S}\exp\left[\frac{\hat{\gamma}_0 - \bar{\gamma}\sigma_{\epsilon_c}^2/r_g}{\bar{\gamma}(1-\sigma_{\epsilon_c}^2/r_g)}\right]. \tag{40}$$

Moreover, the instantaneous BER constraint must be fulfilled at all the rate region boundaries such that

$$\mathrm{BER}(\hat{\gamma}) \leq \mathrm{BER}(\hat{\gamma}_i) = P_b, \\ \hat{\gamma} \in [\hat{\gamma}_i, \hat{\gamma}_{i+1}), \quad 0 \leq i \leq N-1 \tag{41}$$

which by (38) and (22) results in

$$\hat{\gamma}_i = \frac{\bar{\gamma}\sigma_{\epsilon_c}^2}{r_g}\left[1 - \frac{1+A_iS}{A_iS}\ln\left(\frac{P_b}{0.2}(1+A_iS)\right)\right], \\ 0 \leq i \leq N-1. \tag{42}$$

Thus, the cut-off SNR $\hat{\gamma}_0$ and the transmit power S are found through (40) and (42). Once it is done, $\{\hat{\gamma}_i\}_{i=1}^{N-1}$ are easily obtained from (42).

5.3 Average BER and constant power (A-BER, C-Pow)

Finally, we investigate the case concerning the average BER constraint with constant transmit power. Similar to Section 5.2, the transmit power must satisfy (39). The average BER constraint is given by

$$\overline{\mathrm{BER}} \leq P_b, \tag{43}$$

where (38) is used for the instantaneous BER in (25). Forming the Lagrangian function from the criterion (26) and the constraint (43), which is here treated as an equality constraint, gives

$$J(\hat{\gamma}_0, \hat{\gamma}_1, \cdots, \hat{\gamma}_{N-1}) = \sum_{i=0}^{N-1} k_i \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma} \\ +\lambda\left(\sum_{i=0}^{N-1} k_i \int_{\hat{\gamma}_i}^{\hat{\gamma}_{i+1}} \left(\mathrm{BER}(\hat{\gamma}) - P_b\right) f_{\hat{\gamma}}(\hat{\gamma}) d\hat{\gamma}\right). \tag{44}$$

The optimum rate region boundaries are found through solving

$$\frac{\partial J}{\partial \hat{\gamma}_i} = 0, \quad 0 \leq i \leq N-1 \tag{45}$$

which results in

$$\mathrm{BER}(\hat{\gamma}_i) = P_b - \frac{1}{\lambda}, \quad 0 \leq i \leq N-1. \tag{46}$$

Similar to the previous case, we have

$$\hat{\gamma}_i = \frac{\bar{\gamma}\sigma_{\epsilon_c}^2}{r_g}\left[1 - \frac{1+A_iS}{A_iS}\ln\left(\frac{P_b - \frac{1}{\lambda}}{0.2}(1+A_iS)\right)\right], \\ 0 \leq i \leq N-1. \tag{47}$$

We can evaluate the optimal rate region boundaries and transmit power through (40) and (47) based

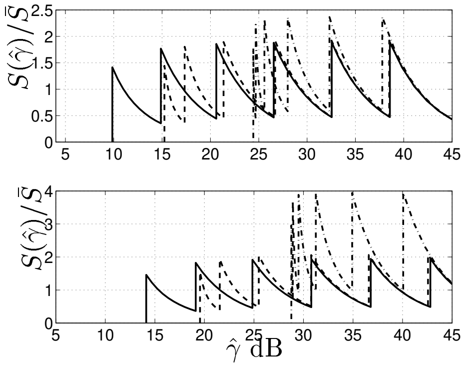

Figure 3: *I-BER, V-Pow* policy: Optimum normalized transmit power versus instantaneous predicted SNR of M-QAM schemes for $\bar{\gamma} = 30$ dB. The upper and lower plots correspond to $P_b = 10^{-3}$ and 10^{-7}, and the solid, dashed and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.001$, 0.01, and 0.1, respectively.

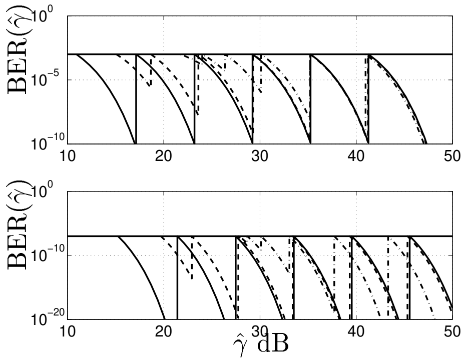

Figure 4: *I-BER, C-Pow* policy: Instantaneous BER versus instantaneous predicted SNR of M-QAM schemes for $\bar{\gamma} = 30$ dB. The upper and lower plots correspond to $P_b = 10^{-3}$ and 10^{-7}, and the solid, dashed and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.001$, 0.01, and 0.1, respectively.

on λ that satisfies the average BER constraint. As shown in [25], the optimal solution that fulfills the average BER constraint with equality exists when $P_b < \frac{1}{1+A_0\bar{S}}$. Otherwise, the solution results in a lower average BER than required with a consequent reduction in the spectral efficiency.

6 Results

We assume that six different M-QAM signal constellations corresponding to 4-QAM, 16-QAM, 64-QAM, 256-QAM, 1024-QAM and 4096-QAM, are available at the transmitter. Also, a flat Rayleigh fading channel with $r_g = 1$ is presumed. The optimal region boundaries for different policies when the required BERs are $P_b = 10^{-3}$ and 10^{-7}, the prediction error variances are $\sigma^2_{\epsilon_p} = 0.001$, 0.01 and 0.1[5] and the average received SNR is $\bar{\gamma} = 30$ dB, can be seen from Figures 3, 4 and 5.

Figure 3 shows that for the *I-BER, V-Pow* policy, the transmit power follows the inverse water-filling pattern w.r.t. $\hat{\gamma}$ within each rate region interval. The peak power within each interval increases as the rate increases. Figure 4 illustrates that under *I-BER, C-Pow*, the instantaneous BER does not exceed the required BER while it reaches the target BER at the boundaries as intended. Finally, the *A-BER, C-Pow* policy results in an instantaneous BER fluctuation around the required average BER to maintain the target BER, on average, as shown in Figure 5.

An interesting phenomenon observed in these figures is the effect of the prediction error variance on the rate region boundaries. We see that when a large prediction error variance is taken into account, the boundaries are *raised* for SNRs lower than the average SNR. To a less extent, they are usually *lowered* for SNRs higher than the average SNR. A reasonable explanation is that when we predict into a fading dip (low SNR), the *relative* conditional prediction standard deviation given by (15), will become larger. This will contribute to making the scheme cautious when entering fading dips. As the prediction error variance is increased, the scheme would not transmit at all during an increasing fraction of the time (when $\hat{\gamma}$ is below $\bar{\gamma}$). Due to the average power constraint, this allows the use of higher transmit power when $\hat{\gamma} >> \bar{\gamma}$. This explains why the SNR limits for use of the largest constellation sizes can be reduced. For more clarification, an example for *I-BER, C-Pow* policy is shown in Figure 6.

The maximum spectral efficiency for $P_b = 10^{-3}$ and 10^{-7}, and $\sigma^2_{\epsilon_p} = 0.001$, 0.01 and 0.1 are il-

[5] For comparison, the use of the average power $r_g = 1$ as a power prediction would result in a prediction error variance $\sigma^2_{\varepsilon_p} = 0.5$ for all prediction horizons. For an average SNR of 20 dB in (2), $\sigma^2_{\varepsilon_p} = 0.1$ corresponds to a prediction 0.4 wavelengths ahead in space, while $\sigma^2_{\varepsilon_p} = 0.01$ is obtained when predicting 0.1 wavelengths ahead when using (7) [21, 22]. The level $\sigma^2_{\varepsilon_p} = 0.001$ is essentially equivalent to perfect prediction of the channel power.

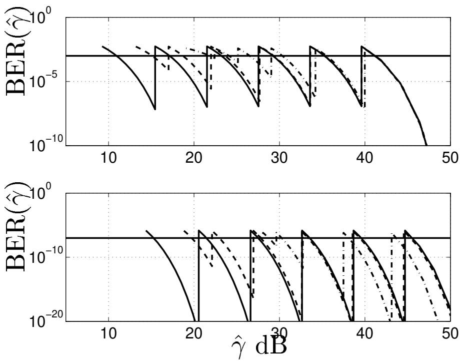

Figure 5: *A-BER, C-Pow* policy: Instantaneous BER versus instantaneous predicted SNR of M-QAM schemes for $\bar{\gamma} = 30$ dB. The upper and lower plots correspond to $P_b = 10^{-3}$ and 10^{-7}, and the solid, dashed and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.001$, 0.01, and 0.1, respectively.

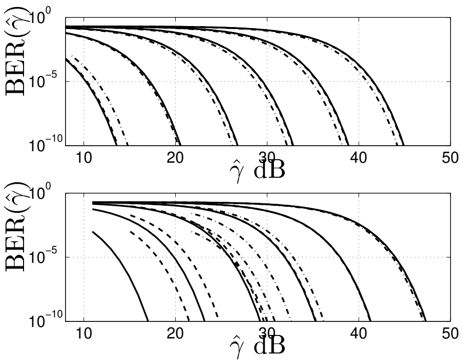

Figure 6: BER versus instantaneous predicted SNR under the *"I-BER, C-Pow"* policy for 4-QAM, 16-QAM, 64-QAM, 256-QAM, 1024-QAM and 4096-QAM schemes. The solid, dashed and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.001$, 0.01 and 0.1, respectively. For a given $\sigma^2_{\epsilon_p}$, the BER curves are moving to the right as the constellation size increases. Moreover, the upper and lower plots correspond to the $\bar{\gamma} = 10$ and 30 dB, respectively.

lustrated in Figures 7, 8 and 9 for *I-BER, V-Pow*, *I-BER, C-Pow* and *A-BER, C-Pow* policies, respectively. The similar trend observed in these figures is that the gain in the spectral efficiency when using good predictors is considerable as compared to the poor predictors. The use of predictors that attain the quality $\sigma^2_{\varepsilon_p} = 0.01$ will result in an insignificant performance deterioration as compared to perfect prediction. For $\sigma^2_{\varepsilon_p} = 0.1$, the reduction in spectral efficiency is rather small at $P_b = 10^{-3}$, but it is large at $P_b = 10^{-7}$.

To facilitate the comparison of different policies from the spectral efficiency point of view, we repeatedly depicted the spectral efficiency results of these schemes for $P_b = 10^{-3}$ and $\sigma^2_{\epsilon_p} = 0.001$ and 0.1, in Figure 10. The results are accompanied with simulation results which seem to be in a good agreement with the numerical ones. We see that for small prediction error variance, the highest and lowest spectral efficiencies are provided by *I-BER, V-Pow* and *I-BER, C-Pow*, respectively. However, as the predictor deteriorates, the spectral efficiency of all the policies become closer to each other. The results in previous figures also show that the difference becomes even negligible when a lower BER is requested.

Finally, Figure 11 is shown to highlight the importance of considering realistic assumptions for the design. In this example, the adaptive modulation systems are designed under the assumption of perfect channel prediction at the transmitter. However, the transmitted signals are experiencing different channels than the predicted ones due to the inaccuracy in the prediction. The results are shown for the channel prediction error variances $\sigma^2_{\epsilon_p} = 0.01$ and 0.1 and target BER $P_b = 10^{-3}$. The two uppermost figures illustrate the instantaneous BER for the average received SNR $\bar{\gamma} = 30$ dB while the lower figure shows the average BER for $\bar{\gamma} = 10$ to 40 dB. It is evident that the target BER will no longer be attained if the prediction errors are neglected in the design of the adaptive modulation scheme.

7 Conclusion

The optimum design of an adaptive modulation scheme based on uncoded M-QAM modulation has been investigated. The transmitter adjusts the transmission rate and possibly also power based on the predicted SNR to maximize the spectral efficiency while satisfying the BER and average transmit power constraints.

Optimum solutions for adjusting the adaptive rate and transmit power have been derived. The analytical results show that when the prediction error increases, the rate region boundaries for a given constellation size are raised for the SNRs lower than average SNR and lowered for the SNRs higher than

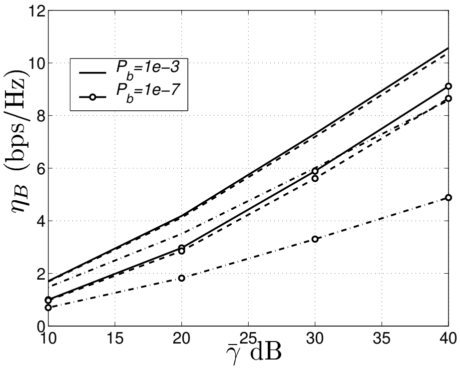

Figure 7: *I-BER, V-Pow* policy: M-QAM Spectral efficiency versus average received SNR for $P_b = 10^{-3}$ and 10^{-7}. The solid, dashed and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.001$, 0.01 and 0.1, respectively.

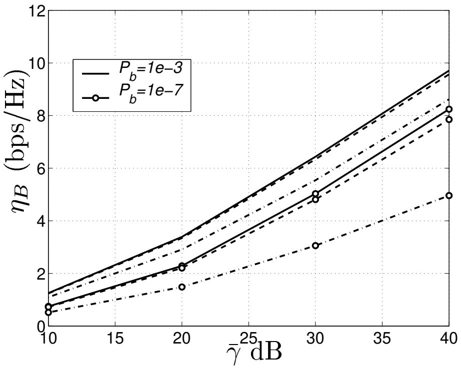

Figure 8: *I-BER, C-Pow* policy: M-QAM Spectral efficiency versus average received SNR for $P_b = 10^{-3}$ and 10^{-7}. The solid, dashed and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.001$, 0.01 and 0.1, respectively.

the average SNR. Moreover, the spectral efficiency decreases as the predictor error variance increases and possibly also the required BER decreases, as expected. Also, the gain due to the transmission with varying power is minor and becomes even negligible when the prediction quality deteriotes. It is demonstrated that the QoS considerably degrades when the system is not design based on realistic assumptions such as erroneous prediction.

The optimization problem that is discussed here, can be solved for any family of modulations, as long as accurate BER expressions (which are invertible and differentiable) in terms of the predicted SNR are available. A competitive candidate is Trellis Coded Modulation (TCM) which will improve the spectral efficiency at a given $\bar{\gamma}$. This is a topic which will be studied in the future.

References

[1] A. J. Goldsmith and P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[2] A. J. Goldsmith, "The Capacity of Downlink Fading Channels with Variable Rate and Power," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 3, pp. 569–580, Aug. 1997.

[3] B. Vucetic, "An Adaptive Coding Scheme for Time-Varying Channels," *IEEE Transactions on Communications*, vol. 39, no. 5, pp. 653–663, May 1991.

[4] S. M. Alamouti and S. Kallel, "Adaptive Trellis-Coded Multiple-Phase-Shift Keying for Rayleigh Fading Channels," *IEEE Transactions on Communications*, vol. 42, no. 6, pp. 2305–2314, June 1994.

[5] M.-S. Alouini and A. J. Goldsmith, "Adaptive M-QAM Modulation over Nakagami Fading Channels," in *IEEE Global Communications Conference*, Phoenix, Arizona, Nov. 1997, pp. 218–223.

[6] A. J. Goldsmith and S. Chua, "Variable-Rate Variable-Power MQAM for Fading Channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[7] V. K. N. Lau and M. D. Macleod, "Variable Rate Adaptive Trellis Coded QAM for High Bandwidth Efficiency Applications in Rayleigh Fading Channels," in *Proc. IEEE Vehicular Technology Conference*, May 1998, vol. 1, pp. 348–352.

[8] A. J. Goldsmith and S. Chua, "Adaptive Coded Modulation for Fading Channels," *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 595–602, May 1998.

[9] T. Ue, S. Sampei, N. Morinaga, and K. Hamaguchi, "Symbol Rate and Modulation Level-Controlled Adaptive Modulation/TDMA/TDD System for High-Bit-Rate Wireless Data Transmission," *IEEE Transactions on Vehicular Technology*, vol. 47, no. 4, pp. 1134–1147, Nov. 1998.

[10] M.-S. Alouini, X. Tang, and A. J. Goldsmith, "An Adaptive Modulation Scheme for Simultaneous Voice and Data Transmission over Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 5, pp. 837–850, May 1999.

[11] Kjell J. Hole, Henrik Holm, and Geir E. Øien, "Adaptive Multidimensional Coded Modulation Over Flat Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1153–1158, July 2000.

[12] C. Köse and D. L. Goeckel, "On Power Adaptation in Adaptive Signaling Systems," *IEEE Transactions on Communications*, vol. 48, no. 11, pp. 1769–1773, Nov. 2000.

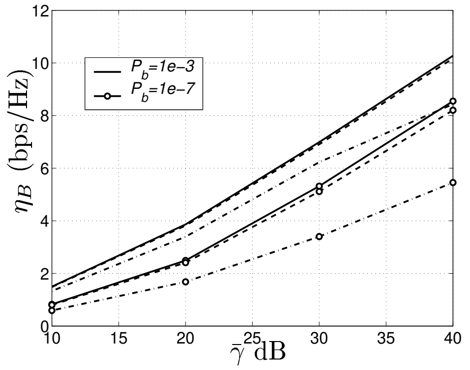

Figure 9: *A-BER, C-Pow* policy: M-QAM Spectral efficiency versus average received SNR for $P_b = 10^{-3}$ and 10^{-7}. The solid, dashed and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.001$, 0.01 and 0.1, respectively.

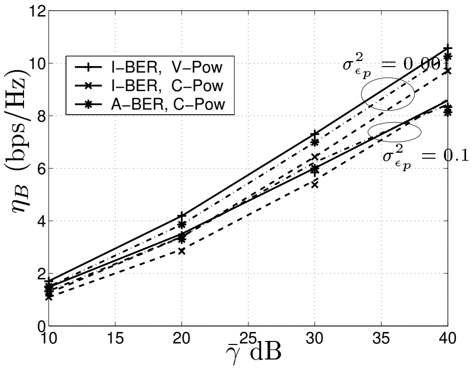

Figure 10: Comparisson of M-QAM spectral efficiency for different policies and $P_b = 10^{-3}$. The solid, dashed and dashed-dotted lines correspond to *"I-BER, V-Pow"*, *"I-BER, C-Pow"* and *"A-BER, C-Pow"* policies, respectively. The markers represent the simulation results.

[13] V. K. N. Lau and M. D. Macleod, "Variable-Rate Trellis Coded QAM for Flat-Fading Channels," *IEEE Transactions on Communications*, vol. 49, no. 9, pp. 1550–1560, Sept. 2001.

[14] S. T. Chung and A. J. Goldsmith, "Degrees of Freedom in Adaptive Modulation: A Unified View," *IEEE Transactions on Communications*, vol. 49, no. 9, pp. 1561–1571, Sept. 2001.

[15] D. L. Goeckel, "Robust Adaptive Coded Modulation for Time-Varying Channels with Delayed Feedback," in *Thirty-Fifth Annual Allerton Conference on Communication, Control, and Computing*, Oct. 1997, pp. 370–379.

[16] D. L. Goeckel, "Adaptive Coding for Time-Varying Channels Using Outdated Fading Estimates," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 844–855, June 1999.

[17] X. Tang, M. Alouini, and A. J. Goldsmith, "Effect of Channel Estimation Error on M-QAM BER Performance in Rayleigh Fading," *IEEE Transactions on Communications*, vol. 47, no. 12, pp. 1856–1864, Dec. 1999.

[18] S. Hu, A. Duel-Hallen, and H. Hallen, "Adaptive Modulation Using Long Range Prediction for Flat Rayleigh Fading Channels," in *IEEE International Symposium on Information Theory*, Sorrento, Italy, June 25-30 2000, vol. 2, p. 159.

[19] S. Hu and A. Duel-Hallen, "Combined Adaptive Modulation and Transmitter Diversity Using Long Range Prediction for Flat Fading Mobile Radio Channels," in *Global Telecommunications Conference*, San Antonio, Texas, November 25-29 2001, vol. 2, pp. 1256–1261.

[20] T. S. Yang and A. Duel-Hallen, "Adaptive Modulation Using Outdated Samples of Another Fading Channel," in *Wireless Communications and Networking Conference*, Orlando, Florida, USA, March 17-21 2002, pp. 477–481.

[21] T. Ekman, M. Sternad, and A. Ahlén, "Unbiased Power Prediction on Broadband Channel," in *Proc. IEEE Vehicular Technology Conference*, Vancouver, Canada, Sept. 2002.

[22] T. Ekman, "Prediction of Mobile Radio Channels, Modeling and Design," Phd thesis, Signals and Systems, Uppsala University, Uppsala, Sweden, Oct. 2002.

[23] M. Sternad, T. Ekman, and A. Ahlén, "Power Prediction on Broadband Channels," in *Proc. IEEE Vehicular Technology Conference*, Rhodes, Greece, May 2001.

[24] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, 4th edition, 2001.

[25] S. Falahati, "Adaptive Modulation and Coding in Wireless Communications with Feedback," Phd thesis, Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, Oct. 2002.

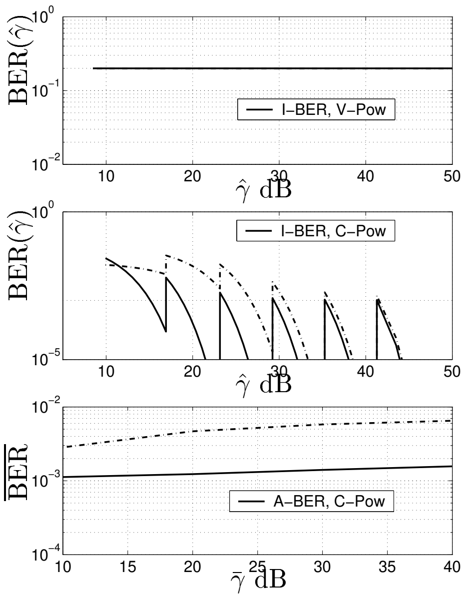

Figure 11: The effect of imperfect CSI on the performace of M-QAM schemes. The results are shown for $P_b = 10^{-3}$. The solid and dashed-dotted lines correspond to $\sigma^2_{\epsilon_p} = 0.01$ and 0.1, respectively.