

Performance Cost of Adaptation in Mobile Wireless Power Control

Minh H. Ha and Robert R. Bitmead

Abstract—Power control in mobile communications is an adaptive control system which functions routinely without direct human intervention. In this paper we examine the cost in terms of power usage of the one-parameter system identification problem of estimating the signal-to-noise ratio and thereby setting subsequent transmissions at the appropriate power level to accommodate the error detection and recovery. This is performed with an explicit formulation of the density function of the signal-to-noise ratio estimate and its dependence on the number of training samples used in this phase - it is a non-central χ^2 density if the noise power is known, else it is F -distributed. The objective is to develop a total cost function based on transmission power under which the price paid for adaptation becomes apparent. Our example is based on GSM mobile telephony systems and is part of a broader study into the costs associated with adaptation.

I. INTRODUCTION

Adaptive systems are dynamic systems which involve an explicit learning or adjustment component in which a number of parameters are adapted based on input-output data. Our aim in this paper is to pose a different aspect of this learning question pertinent to adaptive control and to do this via a simple but familiar example with a single parameter. Yet the principle is applicable to more complicated systems within the context of multi-user and multi-input multi-output setup. We desire to construct a scalar cost function for controlled system performance and to explore the price paid in these system performance terms for the quality of adaptive parameter accuracy. In this manner we hope to analyze the connection between adaptive system behavior and overall closed-loop performance.

Power control is used in mobile wireless communication to preserve battery life as well as manage interference. In this paper, we only consider the battery life problem and treat the optimal transmitted power as that which maintains a fixed pre-specified bit-error-rate (BER). Adaptation involves the online adjustment of transmitted power to attempt to identify the optimal value. But this adaptation comes with a cost in terms of energy consumption. Our aim is to analyze this cost directly in terms of the optimized performance, i.e. power.

The approach adopted is to examine in detail the convergence properties of the Signal-to-Noise Ratio (SNR) estimation algorithm in wireless communications as a function of the number of (*a priori* known and therefore information free) training bits of data sent. The work of Gagliardi

and Thomas [1] employs the same model to estimate the SNR by the maximum likelihood algorithm and assess its validity in terms of the number of received symbols and determine the optimal structure of the estimator based on its ability to evaluate pulse-code modulation (PCM) data reliability. Mohammad and Buehrer [2] also use an SNR estimator to examine the bit-error-rate (BER) performance of adaptive modulation in terms of the SNR estimation error. Our formalism follows initially that of Gagliardi and Thomas [1] in the development of an explicit description of the exact probability density function of the SNR estimate — it is non-central χ^2 or F distributed — in the case of a memoryless fading additive gaussian white-noise (AGWN) channel and a corresponding calculation of the probability of erroneous selection of transmission power. This is then used to study the evolution of this probability as a function of the number of training bits. In turn, this provides an assessment of the relationship between the power used by the training sequence and the concomitant power saving in transmission at the most efficient rate. In this fashion, the adaptation cost of training is measured in system performance units, i.e. energy.

There are a few texts that have a remotely similar theme in a multiuser context, such as [3] where the authors present an approach to estimating SNR using a pilot sequence. They however go no further in applying the algorithm to an adaptive power control context, or quantifying the cost associated with the channel estimation scheme. Goldsmith [4] includes solutions to variable modulation and coding problems and power control based on SNR (or SINR in the case of multiuser networks) but does not formalize the uninformative energy required to carry out the algorithms. Various studies [5], [6], [7] also deal with various types of adaptation problems such as rate, modulation, power and coding adaptations but none has taken into account the tradeoffs in terms of energy. Chi et. al. [8] also aims to design an optimal training sequence in OFDM system with minimum mean square error criterion and examines the tradeoffs that are inherited in OFDM such as peak-to-average power ratio and phase noise. However the paper considers a fixed amount of transmitted energy and is not concerned with finding the optimal amount.

Here our task is more modest in considering a single user problem but takes a step forward to analyze precisely the estimation cost as a penalty on ultimate system performance. The result however can be generalized to a multi-user problem. We shall also see that we further extend this cost to include the cost associated with performing a power control algorithm taking into account the possibility of wrongly selecting an optimal transmitted level, and, in a subsequent

M.H. Ha and R.R. Bitmead are with the Department of Mechanical & Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla CA, 92093-0411, USA. {h5ha, rbitmead}@ucsd.edu. This material is based upon work supported by the US National Science Foundation under Grant No. 1102384.

section, the mobility of the mobile station (MS). This appears to be distinct from other aforementioned studies, including those dealing with only the channel estimation cost and/or the adaptive control and power control problems in general.

II. PROBLEM SET-UP

Consider the transmission of data between a single mobile station (MS) and its serving base station (BS). We make the following assumptions and notational conventions.

- 1) In the stationary MS case, the uplink is described by an AWGN channel with memoryless fading, which we take as fixed, i.e. time-invariant. So the channel is described by a scalar (complex) gain h .
- 2) In the moving MS case, we assume Rayleigh fading. The complex channel gain h now varies.
- 3) The transmission scheme is phase-shift keying (PSK)
- 4) A deterministic sequence of n training data symbols, $\{a_k\}_{k=1}^n$, is sent between the MS and BS at transmitted power P_0 . This data is known to both MS and BS.
- 5) The training data and the noise are independent.
- 6) The MS transmitted training power is $P_0 = \frac{1}{n} \sum_{k=1}^n |a_k|^2$, which is constant with PSK.
- 7) A total of m bits of data (exclusive of the training data) is to be communicated between the MS and the BS.
- 8) The channel noise, $\{w_k\}$, power at the BS receiver is fixed and denoted $\sigma_w^2 = E[w_k w_k^*]$, i.e. $N_0 = \sigma_w^2$.
- 9) Without the training sequence, error detection is used and hence the number of errors is accessible.

III. SNR ESTIMATION & POWER CONTROL WITH TRAINING DATA

A. Bit-error rate & power control

For digital data transmission, the bit-error rate (BER) depends directly on the prevailing SNR in addition to the modulation scheme, as is studied in many texts on communications. For example, for the AGWN channel with binary-PSK (BPSK) the probability of bit-error is given by [9],

$$P_e = Q(\sqrt{2\gamma}) = \frac{1}{2} \operatorname{erfc}(\sqrt{\gamma}). \quad (1)$$

In our context, SNR is equal to the energy per bit to noise power spectral density ratio, E_b/N_0 , at the receiver. This BER curve is plotted on the right vertical scale in Figure 1 and we shall keep this modulation scheme in the subsequent analysis.

For speech communications, a bit error rate of at least 10^{-2} is considered the minimum acceptable [10]. We can then reliably assume in our problem that a reasonable range for BER is from 10^{-4} to 10^{-2} . By (1), the corresponding values of SNR for a BER of 10^{-2} , 10^{-3} and 10^{-4} are 4.32dB, 6.79dB and 8.40dB respectively, which can be seen as roughly 2dB apart. Furthermore, for mobile stations, for example in GSM 1900 (also known as PCS-1900), there are 18 levels of power at which MS can transmit and these are 2dBm apart except for the last two highest levels that are 1dBm apart [11]. For practical purposes, we confine our study to the case that MS uses the transmitted power levels that are uniformly 2dBm apart. Those available transmitted

power levels map into values of the SNR at the BS receiver that are in 2dB steps.

Therefore, the new transmitted power level chosen based upon the SNR estimate will be acceptable if the SNR estimate is within ± 2 dB of the actual SNR at the BS receiver. We shall see shortly that the penalty for the estimate being more than 2dB higher than the actual SNR is different in nature from the penalty for being more than 2dB lower. In the former case, corrupted data are received and the SNR must be relearned. In the latter case, the remaining m bits of data are transmitted at a wasteful power level.

B. SNR estimation with known noise power

Firstly we assume the noise power, σ_w^2 is known. The measurement or observation of the fading channel at the BS has the following form:

$$y_k = ha_k + w_k. \quad (2)$$

The signal power is $S = |h|^2 P_0$ and the SNR is

$$\gamma = \frac{|h|^2 P_0}{\sigma_w^2}. \quad (3)$$

We have access to sequences $\{a_k\}$ and $\{y_k\}$ and σ_w^2 . The aim is to use these n training data bits transmitted at power P_0 to estimate the training SNR at the BS receiver. The BS then communicates to the MS the appropriate increment or decrement in power at which to transmit the remaining m information bits in order that the BER at the BS lie in the range amenable to detection and recovery with the coding scheme in use.

We need an estimate of the SNR using the n most recent measurements $\{y_k\}_{k=1}^n$. To do this, first we will need to evaluate the value of the channel gain h based on these n samples, denoted \hat{h}_n . Adopting the standard least squares channel estimate based on the available data

$$\hat{h}_n = \frac{\frac{1}{n} \sum_{k=1}^n y_k a_k^*}{\frac{1}{n} \sum_{k=1}^n |a_k|^2} = \frac{1}{P_0} \cdot \frac{1}{n} \sum_{k=1}^n y_k a_k^*. \quad (4)$$

Since both $P_0 = \frac{1}{n} \sum_{k=1}^n |a_k|^2$ and σ_w^2 are known at the BS, this suffices for calculation of the empirical signal power via

$$\hat{S}_n = \hat{h}_n \hat{h}_n^* P_0 = |\hat{h}_n|^2 P_0, \quad (5)$$

and SNR estimate

$$\hat{\gamma}_n = \frac{\hat{S}_n}{\sigma_w^2} = \frac{|\hat{h}_n|^2 P_0}{\sigma_w^2} = \left| \frac{1}{n} \sum_{k=1}^n y_k a_k^* \right|^2 \frac{1}{\sigma_w^2 P_0} \quad (6)$$

Using (4), the distribution of \hat{h}_n is $\hat{h}_n \sim N\left(h, \frac{\sigma_w^2}{P_0 n}\right)$
Define

$$X = \frac{\hat{h}_n}{\sqrt{\sigma_w^2/(P_0 n)}}, \quad (7)$$

which is gaussian $N\left(\frac{h}{\sqrt{\sigma_w^2/(P_0 n)}}, 1\right)$. Hence, X^2 is non-centrally chi-square (χ^2) distributed with one degree of freedom and non-centrality parameter $\lambda = \frac{h^2}{\sigma_w^2/(P_0 n)} = n\gamma$.

Now the SNR estimate (6) may be written directly in terms of this $\chi_{1,\lambda}^2$ random variable

$$\hat{\gamma}_n = \frac{1}{n} X^2. \quad (8)$$

Theorem 1: The n -sample SNR estimate, $\hat{\gamma}_n$, with known noise power is distributed as a non-central $\chi_{1,\lambda}^2$ distribution with non-centrality parameter $\lambda = n\gamma$, where n is the number of training bits and γ is the actual SNR. \square

Applying probability density transformation to (8) yields

$$p_{\hat{\gamma}_n}(\hat{\gamma}_n) = np_{X^2}(x^2) \quad (9)$$

It is now possible to plot the pdf of the SNR estimate $\hat{\gamma}_n$. The following is an example demonstrating our calculation. Suppose the parameters are $h = 0.5$, $P_0 = 8\text{dBm}$, or 6mW , which is MS transmitted power level 11 for PCS-1900 [11], and the noise power $\sigma_w^2 = 0.7\text{mW}$. Using (3), the actual SNR is 2.25, or 3.52dB. Figure 1 depicts the pdf of $\hat{\gamma}_n$ for $n = 5, 10, 40$ samples using the left vertical scale. It also shows the BER (1) as a function of SNR for the AGWN BPSK channel using the right vertical scale. Note that the pdf of $\hat{\gamma}_n$ changes with n and γ . Figure 1 indicates that the mean value of the SNR estimate comes closer to the actual SNR as n increases and the variance diminishes.

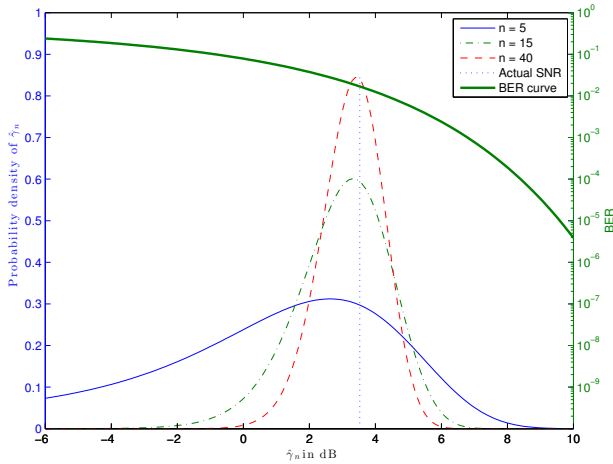


Fig. 1. Probability density of SNR estimate with known σ_w^2 (left vertical scale) and BER versus SNR (right vertical scale).

C. SNR estimation with unknown noise power

When σ_w^2 is not known, the total power – signal plus noise power – needs to be estimated and the noise power estimate extracted. This may be performed by the estimator

$$\hat{N}_n = \frac{1}{n} \sum_{k=1}^n y_k y_k^* - \hat{S}_n. \quad (10)$$

The pdf of this estimate is centrally χ^2 -distributed with $n-1$ degrees of freedom. The SNR estimate is then computed as

$$\hat{\gamma}_n = \frac{\hat{S}_n}{\hat{N}_n} \quad (11)$$

Since $\hat{S}_n \sim \chi_{1,\lambda}^2$ from Theorem 1 and $\hat{N}_n \sim \chi_{n-1}^2$ (central chi-square with $(n-1)$ degrees of freedom), the SNR estimate is then

$$(n-1)\hat{\gamma}_n \sim F(1, n-1; n\gamma),$$

i.e. an F-distribution with $(1, n-1)$ degrees of freedom and noncentrality $n\gamma$.

Theorem 2: The n -sample SNR estimate, $\hat{\gamma}_n$, with the noise power unknown has $(n-1)\hat{\gamma}_n$ distributed as $F(1, n-1; n\gamma)$, the non-central F-distribution with degrees of freedom 1 and $n-1$ and non-centrality parameter $n\gamma$, where n is the number of training bits and γ is the actual SNR. \square

Again we use pdf transformation to get

$$p_{\hat{\gamma}_n}(\hat{\gamma}_n) = (n-1)p_F(f), \quad (12)$$

for $f = (n-1)\hat{\gamma}_n$ and p_F being the probability density function of an $F(1, n-1; n\gamma)$ random variable.

Using the same set of parameters as Section III-B, the pdf of $\hat{\gamma}_n$ now with unknown noise power is plotted in Figure 2.

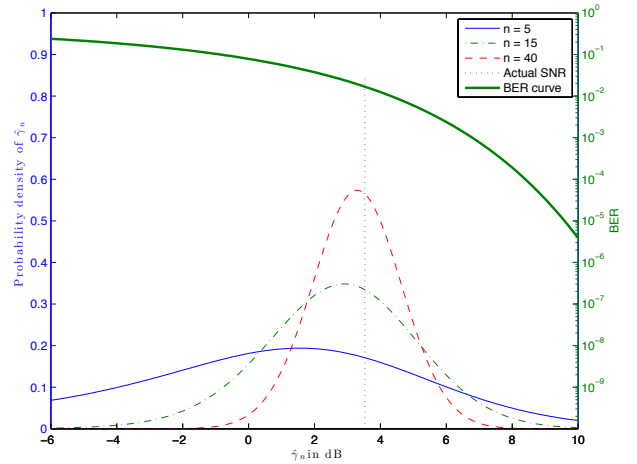


Fig. 2. Pdf of SNR estimate as F-distribution

IV. PRICE OF ADAPTIVE POWER CONTROL

A. Stationary MS

To choose successfully an appropriate MS transmitted power level, there is a price of adaptation in terms of energy spent on transmitting the uninformative training samples. First we consider the case where the MS is stationary. When the MS transmits at a power level higher than needed, the transmission is successful, albeit at a wasteful power. As discussed in Section III-A, the probability of this case happening can be written as a function of n

$$\epsilon_n = P(\hat{\gamma}_n \leq \gamma - 2\text{dB}). \quad (13)$$

The probability that the MS is informed to transmit at a power level lower than the optimal level is

$$\delta_n = P(\hat{\gamma}_n \geq \gamma + 2\text{dB}). \quad (14)$$

In this case, the BS will detect a value of BER higher than acceptable. Assume the BS requires the transmission of l ($l \leq m$) bits to detect that the BER is too high. The best action then is to increase the transmit power, re-estimate the SNR and follow the same power control routine.

Furthermore, as seen in Figure 1, with a large number of training samples n , the probability of $\hat{\gamma}_n$ being two steps (4dB) away from the actual SNR is considerably smaller than the probability of being one step away. Therefore, we assume that if chosen incorrectly, the power level is only one step away from the optimal one. This means $P_1 = P^* + 2dBm$ in the case that $\hat{\gamma}_n \leq \gamma - 2dB$ and $P_1 = P^* - 2dBm$ in the case that $\hat{\gamma}_n \geq \gamma + 2dB$, where P_1 is the power level at which the MS will transmit after the SNR training test and P^* is the optimal power level. Using these ideas, we define the total message energy cost

$$C_n = nP_0 + mP^*P_r + \varepsilon_n m(P^* + 2) + \delta_n(l(P^* - 2) + nP_0 + mP^*P_r + \varepsilon_n m(P^* + 2) + \delta_n(l(P^* - 2) + nP_0 + mP^*P_r + \varepsilon_n m(P^* + 2) + \dots)) \dots).$$

Here:

- nP_0 is the energy cost of the training sequence,
- mP^* is the energy of transmitting the m message bits at the correct/optimal power, P^* ,
- $P_r = 1 - \varepsilon_n - \delta_n$ is the probability of selecting the optimal power P^* after n training samples,
- ε_n is the probability of selecting a transmission power 2dB higher than P^* after n training samples,
- δ_n is the probability of selecting a message transmission power 2dB less than P^* , in which case, after an initial l message samples are sent, the higher BER is detected and the whole process restarted.

We permit unlimited retransmissions of the training sequence leading to an infinite series in δ_n in the cost. Solving the underlying geometric series yields

$$C_n = \frac{nP_0 + mP^*P_r + \varepsilon_n m(P^* + 2) + \delta_n l(P^* - 2)}{1 - \delta_n}. \quad (15)$$

In the case of noise power known (i.e. SNR estimate is χ^2 distributed), using the same parameters in Section III-B and in addition, supposing $m = 1024$ bits and $l = 24$ bits, we have the plot of the cost (15) as a function of the number of training bits for different values of training transmitted power $P_0 = 8dBm$, $10dBm$ and $14dBm$ in Figure 3. Note that because we have not specified a bit rate, the energy cost C_n plotted in Figure 3 is in a nominal scale.

Figure 3 shows three distinct convex curves, each possessing its respective minimum at different values of n . The figure displays the dependence of the learning rate on the training signal power, since the rates of decline of the curves are related to this power as are the initial values of the energy. The energy penalty for additional training bits past the minimum grows as the training power, affecting the depth of the minimum. The minima arrive in the reverse order of training power. For each curve, the informative message energy cost $mP^* = 16.384J$ dominates the training cost. The

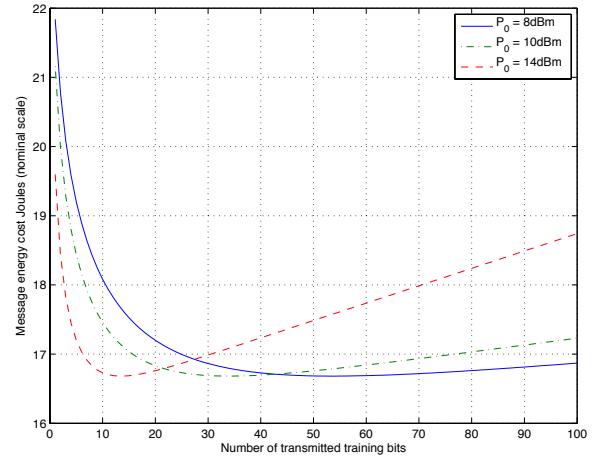


Fig. 3. Total message energy cost with known σ_w^2

minimal cost is achieved by the lowest training power, which naturally exhibits the least sensitivity to the selection of number of training bits.

Next, Figure 4 shows the total message energy cost in the case of unknown noise power (i.e. SNR estimate is F -distributed). The results for the F -distributed case with unknown noise power parallel those for the χ^2 , known noise power case, but differently quantified; the optimizing n is larger for the same P_0 and the training/adaptation cost also is higher, since more parameters must be identified.

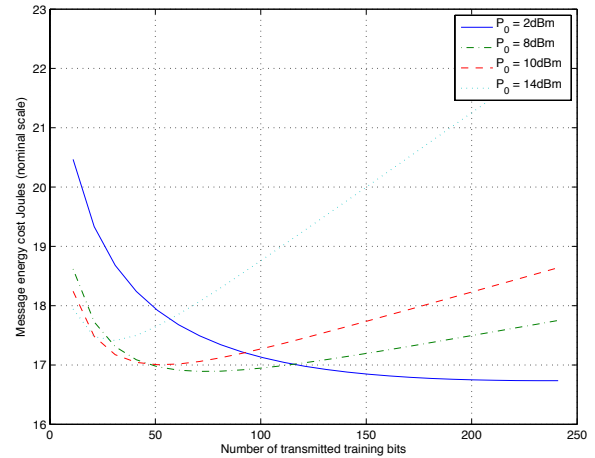


Fig. 4. Total message energy cost with unknown σ_w^2

B. Pathological behavior of the cost function

Figures 3 and 4 indicate that the total message energy cost when the MS is stationary exhibits a pathological behavior. That is, as the training sequence transmission power diminishes, the minimizing value of the training cost, C_n^* , increases. So that, the overall minimizing strategy for a stationary MS is to take training signal power, P_0 , tending

to zero and the optimal signal duration, n^* , tending to infinity. This is a pathological solution, since it relies on the stationarity of the channel and of the optimal power.

To analyze this behavior further in the case of F -distribution, we use a normal approximation [12] of the non-central F -distribution we are dealing with. In [13], we prove that the asymptotic expression for the cost function (15) when n approaches infinity is a function of $\lambda = n\gamma$. Thus, the optimal selection of training power is non-unique and relies on taking longer and longer training sequences of diminishing power, while preserving the total training energy. Clearly, the mobility of the MS must be brought into play to break this pathology.

C. Moving MS

Now we consider the case where the MS is moving with a certain velocity. The Rayleigh fading channel model includes a mobility description, whereby the channel response, h , possesses a correlation function, which is a function of the MS speed expressed as a Doppler shift, f_d

$$R_{hh}(\tau) = J_0(2\pi f_d \tau).$$

Here $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. By analysis in [13], we arrive at the modified cost function

$$\bar{C}_n = [10.75 - 9.75J_0(2\pi f_m n)]C_n. \quad (16)$$

where $f_m = f_d T$ is the maximum Doppler shift f_d normalized by sampling time T . Note that if MS is immobile, we have no Doppler shift, i.e. $f_d = 0$ and $J_0(2\pi f_m n) = 1$. The modified cost function (16) then becomes the original version (15).

We now can plot the modified cost function to see the effect of mobility. For example, take a MS station moving at a speed of 40km/h (common vehicular speed in an urban setup) which corresponds to a maximum Doppler frequency of $f_d = 70.37\text{Hz}$ and a transmission rate of 22.8kbps which is the full-rate speech traffic channel (TCH/FS) bit rate in GSM [14]. Now due to the fast changing nature of the channel, we set $m = 300$. Other parameters remain unchanged.

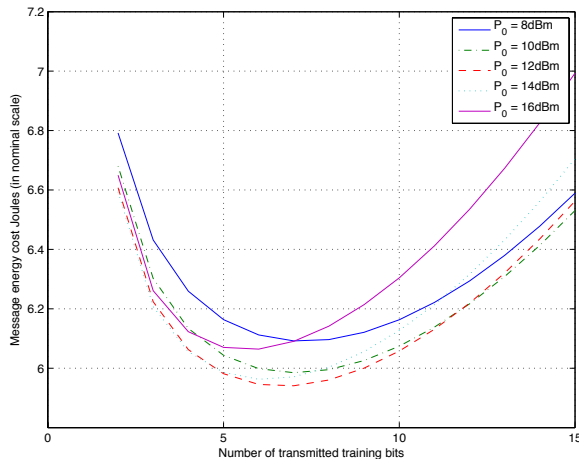


Fig. 5. Modified cost function (MS speed 40km/h)

In Figure 5, we notice that there is an optimal training transmitted power, 12dBm to be exact, that results in an optimal cost (i.e. cost that has the smallest minimum value) across the range of transmitted power levels selected. Note that even at a very low speed, e.g. walking pace, an optimal transmitted power level that is different from the lowest level still exists although it needs more samples but less energy than seen in Figure 5 to reach its minimum.

V. BER-BASED BLIND ADAPTATION

We now consider blind adaptation to avoid the use of the uninformative training sequence and statistical hypothesis testing to detect whether the current BER in the link is lower or higher than a certain usable threshold (for example, 10^{-3}).

Assume that n samples are sent over the link and the number of errors, b , is reported. Random variable b is binomial distributed with parameters (n, p) where n is the sample size and p is the actual BER. The test statistic is the quantity b/n , i.e. the n -sample BER estimate.

The first hypothesis test is to detect whether the BER is higher than 10^{-3} and thus to determine whether or not the transmitted power should be increased.

$$\begin{cases} \text{Null hypothesis} & H_{01} : \text{BER} = p_0 = 10^{-3} \\ \text{Alternative hypothesis} & H_{11} : \text{BER} = p_1 = 10^{-2} \end{cases} \quad (17)$$

The second hypothesis test involves determining whether the BER is lower than 10^{-3} and thus whether or not the transmitted power should be decreased.

$$\begin{cases} \text{Null hypothesis} & H_{02} : \text{BER} = p_0 = 10^{-3} \\ \text{Alternative hypothesis} & H_{12} : \text{BER} = p_2 = 10^{-4} \end{cases} \quad (18)$$

Following convention, denote by α the probability of Type-I error, rejection of the null hypothesis when it is true, set to 5% here. Figure 6 depicts these 95% confidence values for the two hypothesis tests. In the figure, if b/n , exceeds the “x” values for n samples, we reject the null hypothesis and declare the BER too high. Likewise, if b/n falls below the “o” values for a given n , we declare the BER too low.

To moderate this decision, the power, $1 - \beta$, of these tests is evaluated by computing the corresponding probability, β , of a Type-II error, i.e. accepting the null hypothesis when it is false. Figures 7 and 8 display the power of each test as a function of sample size, n . Traditionally a power value of 80% is chosen [15]. Figures 6 (“x” values) and 7 show that for the decision to be made that the BER is too high with confidence 95% and power 80% one would require seeing more than one error in a sample of 350 bits, or, more than two errors in 360 to 810 samples. Similarly, Figures 6 (“o” values) and 8 indicate that we would need to see less than one error in 3000 samples to decide confidently that the BER was too low.

From the perspective of mobile Rayleigh fading channels, it is apparent that these sample lengths, 350 and 3000, are too large to accommodate reasonable mobilities. That is, training is necessary with realistic mobilities.

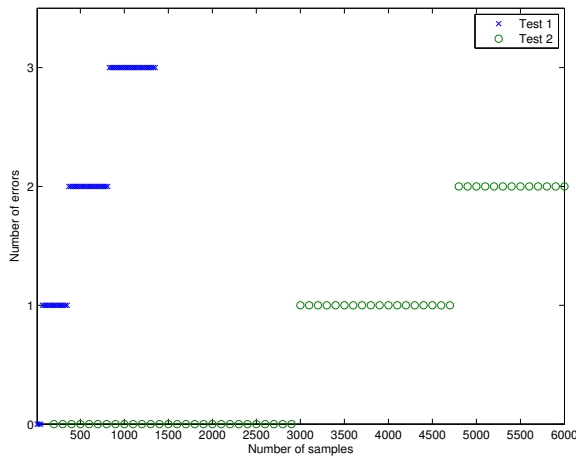


Fig. 6. Number of errors b in n samples to reject the null hypothesis with 95% confidence, above the blue values (BER too high) or below the green values (BER too low).

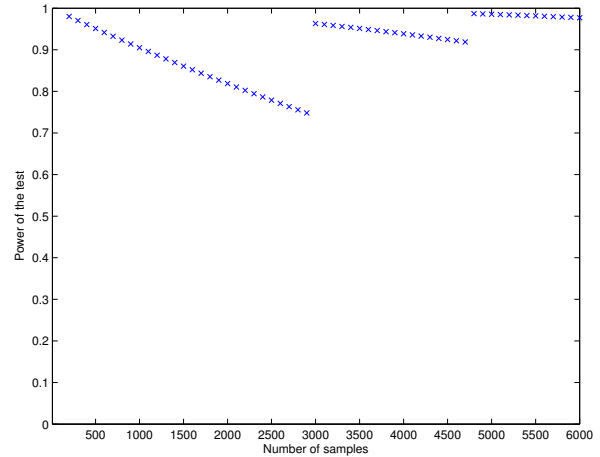


Fig. 8. Power values for the second hypothesis test: BER too low.

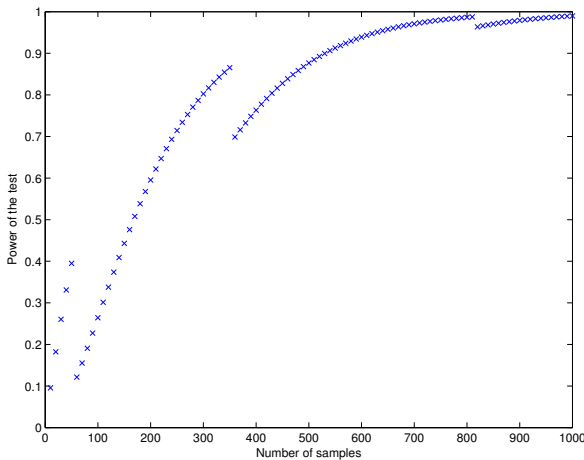


Fig. 7. Power values for the first hypothesis test: BER too high.

VI. CONCLUSIONS

The cost of adaptation based on a training sequence is formulated for both the stationary MS case, where it exhibits a pathological behavior, and the moving MS case, where an optimal transmitted power exists but depends on various factors including the Doppler shift introduced by Rayleigh fading in an urban environment. Without training sequences, BER-based blind adaptation is developed and shown to be inappropriate for power control because of its long latency. The approach taken here is new, although the problem domain is well worn. That is, we consider the cost of adaptation directly in terms of the quantity, power, which adaptation seeks to optimize. We have developed a specific energy cost function and shown that it has a minimal value for certain training sequence lengths and powers, provided mobility is included as part of the formulation. We have also demonstrated that training-sequence-free methods of blind

adaptation are incapable of managing the time-variability of Rayleigh fading channels. In future work [13], we shall quantify the least value of the cost of including adaptation into this problem.

REFERENCES

- [1] R. Gagliardi and C. Thomas, "PCM data reliability monitoring through estimation of Signal-to-Noise Ratio," *IEEE Trans. Communication Technology*, vol. COM-16, no. 3, pp. 479–486, June 1968.
- [2] M. Mohammad and R. M. Buehrer, "On the impact of SNR estimation error on adaptive modulation," *IEEE Comm. Letters*, vol. 9, no. 6, pp. 490–492, June 2005.
- [3] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge Univ. Press, 2005.
- [4] A. Goldsmith, *Wireless communications*. Cambridge Univ. Press, 2005.
- [5] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. on Comm.*, vol. 47, no. 6, pp. 844–855, June 1999.
- [6] S. G. Chua and A. Goldsmith, "Variable-rate variable-power MQAM for fading channels," *IEEE 46th Veh. Tech. Conf.*, vol. 2, pp. 815–819, 1996.
- [7] K. Balachandran, S. R. Kadaba, and S. Nanda, "Channel quality estimation and rate adaptation for cellular mobile radio," *IEEE JSAC*, vol. 17, no. 7, pp. 1244–1256, July 1999.
- [8] Y. Chi, A. Gomaa, N. Al-Dhahir, and A. R. Calderbank, "Training sequence design and tradeoffs for spectrally-efficient multi-user MIMO-OFDM systems," *IEEE Trans. on Wireless Comm.*, vol. 10, no. 7, pp. 2234–2245, July 2011.
- [9] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, 2nd ed. The Netherlands: Kluwer Academic Publishers, 1990.
- [10] A. Mehrotra, *GSM System Engineering*, ser. Mobile Communications. Norwood, MA: Artech House, Inc., 1997.
- [11] ETSI, *Digital cellular telecommunications system (phase 2+); Radio transmission and reception (3GPP TS45.005)*, v10.3.0 release 10 ed., ETSI, France, Jan. 2012.
- [12] N. C. Severo and M. Zelen, "Normal approximation to the chi-square and non-central F probability functions," *Biometrika*, vol. 47, no. 3 and 4, pp. 411–416, 1960.
- [13] M. H. Ha and R. R. Bitmead, "Performance cost of adaptation in mobile wireless power control," in preparation.
- [14] M. Saily, G. Sebire, and E. Riddington, *GSM/EDGE evolution and performance*. Wiley, 2011.
- [15] J. P. M. de Sa, *Applied statistics using SPSS, STATISTICA, MATLAB and R*. Springer, 2007.