Adaptive Coding with Pilot Signals.

Muriel Médard[1], Ibrahim Abou-Faycal[1], Upamanyu Madhow[2]

[1] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology

[2] Electrical and Computer Engineering Department, University of California Santa Barbara

Abstract

We consider adaptive modulation and coding for channels with pilot symbol assisted modulation. Unlike most adaptive transmission schemes, we do not have any channel side information at the sender. The sender adapts not to the state of the physical channel traversed by its transmitted signal, but to the quality of the measurement afforded by pilot symbols at the receiver. We consider the case where no pilot symbols are used, where pilot symbols are used with adaptive reception but without adaptive transmission, and the case where both reception and transmission are done adaptively (while maintaining the average per symbol energy constant). Our results are in terms of capacity for binary signaling. We show that, under appropriate optimizing conditions for the spacing between consecutive pilot symbols, pilot symbol assisted modulation is helpful in terms of capacity. Moreover, adaptation at both the sender and the receiver yields clear advantages over adaptation at the receiver only.

In order to make best use of wireless resources, many different types of adaptive schemes are employed. Adaptive schemes seek to modify the transmission scheme used by the sender according to the state of the channel seen by the receiver. Generally, such schemes involve feedback, concerning the state of the channel, from the receiver to the sender. In an information-theoretic context, adaptive signaling is used for Markov channels with perfect sender and receiver channel side information [Wol78, GV97] or imperfect channel side information [CS99, MS00, Vis99, Kle01]. Power control is a commonly used type of adaptive transmission. Many practical schemes consider modifying the modulation used in order to combat fading. A common means of adapting transmission is to use different types of modulation [TH99, QC99], for instance different levels of QAM constellations [WH00, WS95, SKM94, TH97, GC97, LJ98, WYH00, THK99], according to the state of the channel and possibly other considerations such as multi-user interference [GA99].

In this paper, we consider a different type of adaptive scheme. While the scheme still adapts the transmission to the channel seen by the receiver, the transmitter does not use feedback to determine its policy. Instead, the transmitter takes into account the time-varying quality of the channel measurement available at the receiver in order to modify its transmission policy. Thus, the transmitter adapts to the quality of the channel measurement, rather than to the quality of the channel (in terms of carrier to noise ratio or other metric.)

Our model is the following. We consider a single sender and receiver, connected by a time-varying Rayleigh fading channel. The Rayleigh fading channel is modeled as a Gauss-Markov process. The sender transmits coded and modulated data. The sender has no information regarding the state of the channel and, therefore, does not adapt its transmission scheme in response to fades. Moreover, at regular intervals, the sender transmits a constant and known pilot symbol, whose purpose is to enable measurement

terrupted by pilot symbols. Such a scheme is often referred to as pilot symbol assisted modulation [TAG99, KKJ+97, Cav91, SS93, TH95]. Pilot symbols are commonly used to improve detection [HK96, MB84] and decoding [WC99]. The pilot symbols have energy equal to the average energy constraint. The only channel estimate available at the receiver comes from the pilot symbols, hence there is no data-directed estimation of the channel. The channel estimate at a time sample k depends on the received pilot symbols through its position with respect to the pilot symbols. Given the estimate of the channel obtained through the pilot symbols, the channel is no longer Rayleigh but Ricean with a known specular component. We consider binary signaling, since such signaling performs well at low SNRs and achieves capacity for low SNRs for Rayleigh channels. We also consider a discrete time, sampled system.

Our purpose is to asses the benefits of performing adaptive transmission which takes into account the quality of the estimate obtained at the receiver. We seek to maximize mutual information. The maximization of mutual information subsumes the optimization of both the modulation and the coding. We consider three cases. First, we consider the case where no pilot symbols are transmitted and the channel is not estimated at the receiver. The channel at the receiver is then a Rayleigh channel. Since we do not perform any data-directed estimation, the capacity of the system is the same as if all the channel samples were mutually independent (by averaging over very long times). The capacity of such channels was found in [AFTS97]. For low to moderate SNRs, binary signaling is optimal. The purpose of considering the Rayleigh channel with no pilot symbols is to establish a basis of comparison for the other two cases, which do employ pilot symbols. Indeed, the first question we pose is whether it is preferable to forego pilot symbols and channel estimation altogether and devote to coded data the time allocated to pilot symbols.

Second, we consider the case where we use pilot symbols to aid in the detection and decoding at the receiver but do not modify the distribution of the transmitted signal. We term this scheme the non-adaptive scheme with pilot tones, since the scheme is adaptive at the receiver but not at the sender. The sender uses the distribution that is optimal for transmission when the channel is block-faded. In effect, the sender behaves as though the channel estimate did not vary between pilot symbols. Since we look at mutual information, we may consider this case to be an upper bound to the case where neither the distribution nor the coding is modified according to the distance between the transmitted signal and the pilot symbols. Alternatively, we may view our results as representing capacity when the distribution of the transmitted signal is fixed.

Third, we consider the case where we use pilot symbols and adaptive transmission, but maintain the average per symbol power constant in the coded data. We consider the case where the estimation of the channel is causally performed using the last pilot symbol, and the case where the estimation is non-causally based on the last transmitted pilot along with the pilot transmitted next. At each time sample, the sender changes his transmission policy according to the distance to pilot symbols. We compare the performance of the second and third cases in order to establish the benefit of adaptive transmission.

In Section 2, we present our channel model and the derivations for both the non-adaptive and adaptive coding with pilot symbols. In Sections 2 and 3, we discuss in more detail the non-adaptive and adaptive schemes used respectively and in Section 4, we present our numerical results. These results allow us to optimize numerically the spacing between pilot symbols for adaptive and non-adaptive schemes, since such an optimization cannot be obtained in closed form. Using the optimized spacing between pilot symbols for the different schemes, we can evaluate the benefit of adaptive schemes.

1 Channel Model

We consider the following discrete-time model for the Rayleigh fading channel

$$Y = AX + N,$$

where X is the channel input, Y the output, and A and N are independent complex circular Gaussian random variables with mean zero and variance σ_A^2 and σ_N^2 respectively. Equivalently, the amplitude of the fading coefficient A is Rayleigh distributed and its phase is uniform. The input X is average power limited: $E\left[|X|^2\right] \leq P$.

We assume that the fading process can be represented as a first-order Gauss Markov process

$$A_k = \alpha A_{k-1} + Z_{k-1},$$

where k denotes the time index. Depending on the scale of the time variations of the channel, α^1 may vary between one (corresponding to no channel variations) and zero (where the fading is fast, changing independently from one symbol time to the next.) A similar channel model was introduced in [M00], where the relation between this channel and the coherence time is established.

For a coherence time of T_c and transmission over a bandwidth of W, α is determined by

$$\alpha^{T_c W} = \phi \tag{1}$$

where ϕ is the level of decorrelation we deem necessary in our definition of coherence time. In the literature, the correlation coefficient is taken to vary from 0.9 ([CL75]) to 0.37 ([BN63]) for a time separation of T_c. For bandwidths in the 10 kHz range, and Doppler spreads of the order of 100 Hz, α will typically range between 0.9 and 0.99. For instance, for a Doppler spread of 100 Hz, $W = 10^4$ Hz, $\phi = 0.1$, we have $\alpha = 0.977$. For a Doppler spread of 200 Hz, $\alpha = 0.955$. For a Doppler spread of 50 Hz, $\alpha = 0.989$.

When no pilot signals are used and the correlation between fading coefficients is ignored, the channel behaves as a memoryless Rayleigh fading channel. The capacity of this channel was studied in [AFTS97] and the optimizing input distribution was found to be discrete with a finite number of mass points one of then located at the origin. Furthermore, a binary distribution was found to be optimal at low and moderate values of the SNR motivating the use of binary input strategies in this paper.

Here we consider the case where a pilot signal of power P is transmitted once at the beginning of each interval of length T, enabling the receiver to estimate the fading coefficients governing the channel statistics. The resulting channel outputs can be written as

$$Y_{lT} = A_{lT}\sqrt{P} + N_{lT}, \quad l \in \mathbb{N}.$$

Based on the observation of $\{y_{lT}\}_l$, estimates of the fading coefficients $\{A_k\}_{lT+1}^{(l+1)T-1}$ are obtained and at each time step k, using its estimate the receiver experiences a Ricean

---

[1]Since the fading is Rayleigh distributed, it is sufficient to consider only real positive values for the parameter α.

estimate of the value of A_k and v_k is the variance of the estimation error:

$$p_{Y_k|X_k}(y_k|x_k) = \frac{1}{\pi(v_k|x_k|^2 + \sigma_N^2)} \exp\left\{-\frac{|y_k - \hat{A}_k x_k|^2}{v_k|x_k|^2 + \sigma_N^2}\right\}, \quad lT < k < (l+1)T.$$

Therefore[2], for a given received sequence $\{y_{lT}\}_l$ the mutual information $I_k(X_k; Y_k|\{Y_{lT}\} = \{y_{lT}\})$ at time $k \in \{lT+1, \cdots, (l+1)T-1\}$ is given by

$$\begin{aligned}
I_k\big(X_k; Y_k|\{Y_{lT}\} = \{y_{lT}\}\big) = & p_k(1) \int p_{Y_k|X_k}(y|x_k(1)) \ln \frac{p_{Y_k|X_k}(y|x_k(1))}{p_{Y_k}(y)} dy \\
& + (1 - p_k(1)) \int p_{Y_k|X_k}(y|x_k(2)) \ln \frac{p_{Y_k|X_k}(y|x_k(2))}{p_{Y_k}(y)} dy,
\end{aligned}$$

where the probability mass function of the binary input $\{x_k(1), x_k(2)\}$ is $\{p_k(1), (1 - p_k(1))\}$.

Furthermore, since we are eliminating the possibility of data-directed estimation, the achievable rates[3] are hence given by

$$\mathsf{E}\left[\frac{1}{T} \sum_{k=lT+1}^{(l+1)T-1} I_k\big(X_k; Y_k|\{Y_{lT}\} = \{y_{lT}\}\big)\right] = \frac{1}{T} \sum_{k=lT+1}^{(l+1)T-1} \mathsf{E}\left[I_k\big(X_k; Y_k|\{Y_{lT}\} = \{y_{lT}\}\big)\right],$$

$$(2)$$

where the expectation is over the (correlated) random variables $\{Y_{lT}\}$ which are circular Gaussian distributed with mean zero and variance $\sigma_A^2 P + \sigma_N^2$.

Overall we consider two estimation procedures:

1. The receiver performs a causal estimation of the channel parameters based on the most recently sent pilot tone. For each interval, based on the observation of y_{lT}, estimates of the fading coefficients $\{A_k\}_{lT+1}^{(l+1)T-1}$ are obtained by a standard Bayesian least square estimation procedure yielding

$$\hat{A}_k = \frac{\sqrt{P}\sigma_A^2}{P\sigma_A^2 + \sigma_N^2} \alpha^{(k-lT)} y_{lT} \tag{3}$$

$$v_k = \sigma_A^2 - \frac{P\sigma_A^4}{P\sigma_A^2 + \sigma_N^2} |\alpha|^{2(k-lT)}, \quad lT < k < (l+1)T. \tag{4}$$

2. The estimation is non-causal and based on the most recently sent pilot tone along with the one transmitted next, i.e. $\hat{A}_k$ is based on the values of y_{lT} and $y_{(l+1)T}$ whenever $lT < k < (l+1)T$. In this case

$$\begin{aligned}
\hat{A}_k = & \frac{\sqrt{P}\sigma_A^2}{(P\sigma_A^2 + \sigma_N^2)^2 - P^2\alpha^{2T}} \left[\left[\alpha^{(k-lT)}(P\sigma_A^2 + \sigma_N^2) - P\alpha^{2T-(k-lT)}\right] y_{lT} \right. \\
& \left. + \left[\alpha^{T-(k-lT)}(P\sigma_A^2 + \sigma_N^2) - P\alpha^{T+(k-lT)}\right] y_{(l+1)T}\right]
\end{aligned}$$

$$(5)$$

$$v_k = \sigma_A^2 - \frac{P\sigma_A^4}{(P\sigma_A^2 + \sigma_N^2)^2 - P^2\alpha^{2T}} \left[(P\sigma_A^2 + \sigma_N^2)\left[\alpha^{2(k-lT)} + \alpha^{2[T-(k-lT)]}\right] - 2P\alpha^{2T}\right]$$

$$(6)$$

---

[2] Under the assumption of binary signaling.
[3] Note that the mutual information at times $k = lT$ is zero.

4

First we consider the scheme where the transmitter does not adapt its transmission strategy to the statistics of the channel estimates used at the receiver. More precisely, we compute the achievable rates when the transmitter is using a single fixed input distribution at all times. This of course will allows us to quantify the performance of a system when the transmitter is not adapting its coding strategy and using instead one fixed codebook independently of the time index.

We look mainly at the case where the transmitter considers the channel to be block-faded where, the fading coefficient is assumed constant over intervals of length T and changing independently from one interval to the next. Consequently, we find first the optimal input distribution for the block-faded system, and then compute the average mutual information under this distribution for different values of α. While when $\alpha = 1$ the model based on the first estimation procedure would correspond to block-fading, as α decreases we expect the performance of this non-adaptive system to deteriorate, and we quantify this performance loss.

3   Adaptively Coded System

Here we consider the scheme where, without any channel state information, the transmitter takes into consideration the statistics of the channel estimates at the receiver. It adapts accordingly its modulation and coding to maximize the rates that can be reliably transmitted over the channel. While no optimal power allocation is performed here (a constant amount of power is used instead,) at each time step the transmitter uses a "good" codebook achieving the highest mutual information of the Ricean channel the receiver sees.

Equivalently, one can think of the problem as that of finding the best input strategy that maximizes the expected mutual information $\mathsf{E}\left[I_k\big(X_k; Y_k | \{Y_{lT}\} = \{y_{lT}\}\big)\right]$ for each time step between $lT + 1$ and $(l + 1)T - 1$. For these computations, we considered the two estimation methods described in section 1, i.e. for the pair $\left(\{\hat{A}_k\}, \{v_k\}\right)$ given by either equations (3) and (4), or equations (5) and (6).

We use standard matlab tools to optimize, for each time period $k \in \{1, \cdots, T-1\}$, the expected mutual information over the input probability distribution. The corresponding optimal distribution yields of course the highest achievable rates depending on how far the transmission is occurring with respect to the pilot signals.

Sending pilot tones frequently clearly reduced the rates as a significant portion of the time and power is used to estimate the channel and no information is conveyed from the transmitter to the receiver. On the other hand, when the pilots are used very infrequently, the channel estimates at the receiver are of poor quality and the information rates are low. One of the questions that we will try to answer in the following section is: what is the optimal value of T that would yield the best compromise?

4   Numerical Results

In figures 2 and 3 we have drawn the achievable rates for different values of the SNR, for $\alpha = 0.95$ and 0.99 respectively. In these plots the value of T varies between 2 and 30.

The dashed lines in these plots represent the performance of the system described in section 2, while the solid lines are those of the adaptively coded system of section 3, when the causal estimation method is used.

Figure 1: Best achievable rates function of T and the SNR for $\alpha = 0.9$
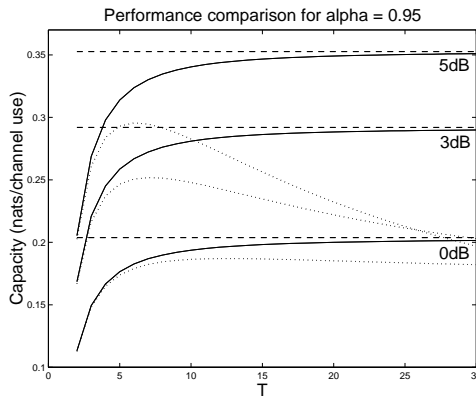

Figure 2: Best achievable rates function of T and the SNR for $\alpha = 0.95$

As it is apparent from figure 2, in the considered ranges of the SNR, there is no benefit gained from using pilot symbol assisted modulation for $\alpha = 0.95$. Indeed, for these low correlation models, not sending a pilot tone and coding instead to a memoryless Rayleigh fading channel outperforms the adaptive and non-adaptive causal scheme we have described.

However, for $\alpha = 0.99$, figure 3 not only shows that an improvement is possible, but also that there is a compromise between infrequent transmission of pilots, and channel estimate's quality. Indeed, for every value of the SNR there is an optimal value for T where the rates are maximized, and these optimal values are apparent in the figure. This holds for both, the adaptive and non-adaptive schemes and the relative improvement that can be achieved with the adaptive scheme considered ranges between 4.2% and 7.5% depending on the SNR. Moreover, as expected, the adaptive scheme outperforms the non-adaptive one by 1.5% to 3.2% when the SNR ranges between 0 and 5dB. This gain becomes more significant for smaller values of α as it can be seen in figure 2.

In figures 4 and 5, we have drawn the achievable rates for the same values of the SNR, α and T, when the non-causal estimation method is used.

While it is clear that, as expected, this scheme outperforms the previous one, the optimal values of T appear to be smaller. Initially, one would expect the opposite since the quality of the channel estimates for large values of T is significantly better especially in the mid-interval range. This suggests that choosing a large value of T is beneficial, especially
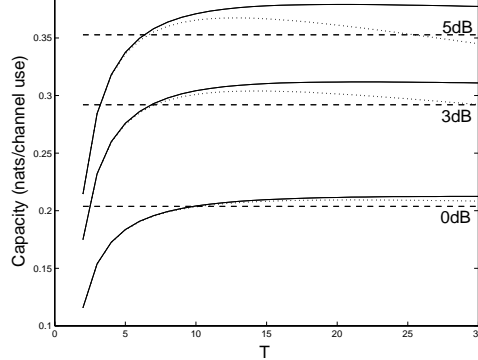
Figure 3: Best achievable rates function of T and the SNR for $\alpha = 0.99$
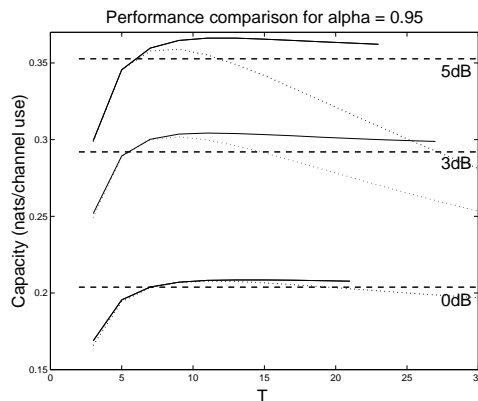

Figure 4: Best achievable rates function of T and the SNR for $\alpha = 0.95$

that it reduces the fraction of time when the channel is idle. However, this is not the case. Intuitively speaking, the gain one gets from the two-pilots based channel estimation is considerable making the more frequent transmission of pilots here significantly better.

For illustrative purposes, in figures 6, 7 and 8 we have drawn the achievable rates for given values of the SNR, as a function of α when T varies between 2 and 30.

Similarly, in figure 9 we have drawn the achievable rates when the two-pilots method is used.

5   Conclusions and Future Work

We have investigated the use of adaptive and non-adaptive sender strategies for time-varying channels with pilot symbol assisted modulation and no feedback from the receiver to the sender. We have shown that, depending on the rate of change of the channel, adaptive sender strategies, properly optimized for spacing between consecutive pilot tones, can improve channel capacity.

Our adaptive strategy maintains average power constant on a symbol-per-symbol basis. A natural extension is to investigate to what extent adaptive power modulation can further improve capacity. Establishing an optimal power allocation is difficult, however, because we do not have a closed form for capacity. Preliminary results on some heuristics methods, akin to waterfilling, which attempt to maintain the interference constant, do not generally lead to improvements over constant power allocations.
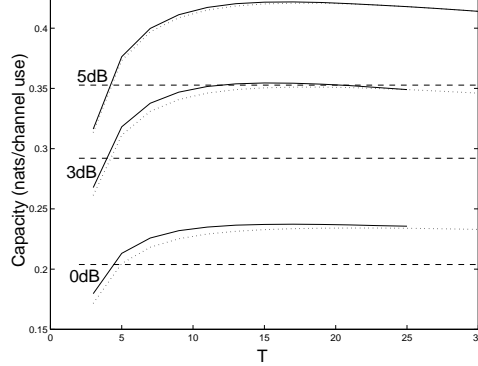
Figure 5: Best achievable rates function of T and the SNR for $\alpha = 0.99$
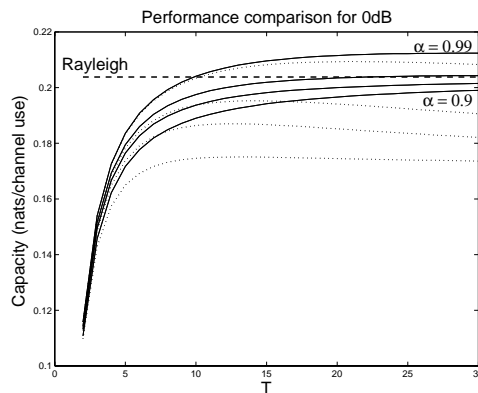

Figure 6: Best achievable rates at 0dB function of T when $\alpha = 0.9, 0.95, 0.97$ and 0.99

References

[AFTS97] I.C. Abou-Faycal, M.D. Trott, and S. Shamai (Shitz). The capacity of discrete-time Rayleigh fading channels. *ISIT*, 8(2):473, 1997.

[BN63] P.A. Bello and B.D. Nelin. The effect of frequency selective fading on the binary error probabilities of incoherent and differentially coherent matched filter receivers. *IEEE Transactions*, CS-11(2), June 1963.

[Cav91] J.K. Cavers. An analysis of pilot symbol assisted modulation for Rayleigh fading channels. *IEEE Transactions on Vehicular Technology*, 40:686–693, November 1991.

[CL75] D.C. Cox and R.P. Leck. Correlation bandwidth and delay spread multipath propagation statistics for 910 mhz urban mobile radio. *IEEE Transactions*, COM-23(1), November 1975.

[CS99] G. Caire and S. Shamai. On the capacity of some channels with channel state information. *IEEE Transactions on Information Theory*, 45(6):2007–2019, September 1999.

[GA99] A.J. Goldsmith and M.-S. Alouini. Area spectral efficiency of cellular mobile radio systems. *IEEE Transactions on Vehicular Technology*, 48(4):1047–1066, July 1999.
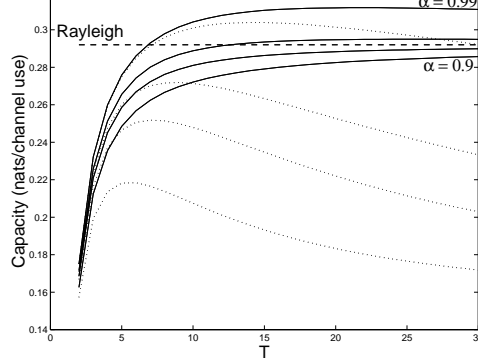
Figure 7: Best achievable rates at 3dB function of T when $\alpha = 0.9, 0.95, 0.97$ and 0.99
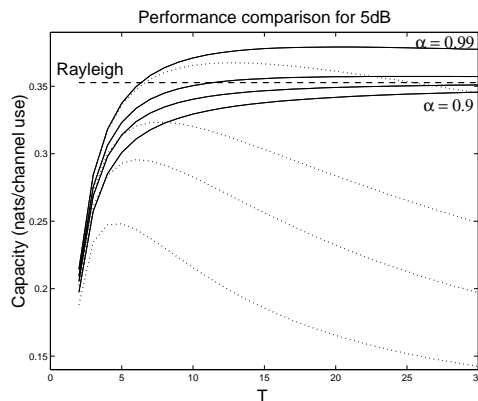

Figure 8: Best achievable rates at 5dB function of T when $\alpha = 0.9, 0.95, 0.97$ and 0.99

[GC97]   A. Goldsmith and S.G. Chua. Variable rate variable power MQAM for fading channels. *IEEE Transactions on Communications*, 45:1218–1230, October 1997.

[GV97]   A.J. Goldsmith and P. Varaiya. Capacity of fading channels with channel side information. *IEEE Transactions on Information Theory*, November 1997.

[HK96]   P. Ho and J.H. Kim. Pilot symbol-assisted detection of cpm schemes operating in fast fading channels. *IEEE Transactions on Communications*, 44:337–347, March 1996.

[KKJ+97] Y.S. Kim, C.J. Kim, G.Y. Jeong, Y.J. Bang, H.K. Park, and S.S. Choi. New Rayleigh fading channel estimator based on PSAM channel sounding technique. *icc*, pages 1518–1520, June 1997.

[Kle01]   T.E. Klein. *Capacity of Gaussian Noise Channels with Side Information and Feedback*. MIT Ph.D. Thesis, February 2001.

[LJ98]   C.H. Lim and J.K. Jeong. Adaptive modulation using the multipath fading compensation. *IEEE Transactions on Vehicular Technology*, 34:940–942, 1998.

[Mé00]   M. Médard. The effect upon channel capacity in wireless communication-sof perfect and imperfect knowledge of the channel. *IEEE Transactions on Information Theory*, 46(3):935–946, 2000.
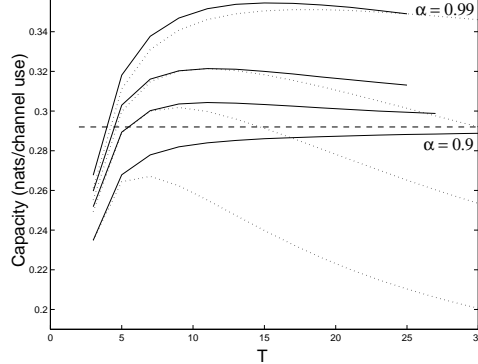
Figure 9: Best achievable rates at 3dB function of T when $\alpha = 0.9, 0.95, 0.97$ and 0.99

[MB84]   J.P. McGeehan and A.J. Bateman. Phase locked transparent tone-in-band (TTIB): a new spectrum configuration particularly suited to the transmission of data over SSB mobile radio network. *IEEE Transactions on Communications*, COM-32:81–87, January 1984.

[MS00]   M. Médard and R. Srikant. Capacity of nearly-decomposable markovian fading channels under asymmetric receiver-sender side information. In *International Symposium on Information Theory*, page 413, 2000.

[QC99]   X. Qiu and K. Chawla. On the performance of adaptive modulation in cellular systems. *IEEE Transactions on Vehicular Technology*, 48(4):1237–1251, July 1999.

[SKM94]   S. Sampei, S. Komaki, and N. Morinaga. Adaptive modulation/ TDMA scheme for large capapcity personal multimedia communications systems. *IEICE Transactions on COmmunications*, E77-B:1096–1103, September 1994.

[SS93]   S. Sampei and T. Sunaga. Rayleigh fading compensation for qam in land mobile radio communications. *IEEE Transactions on Vehicular Technology*, 42:137–147, May 1993.

[TAG99]   X. Tang, M.-S. Alouini, and A.J. Goldsmith. Effect of channel estimation error on M-QAM BER performance in Rayleigh fading. *IEEE Transactions on Communications*, 47(12):1856–1864, December 1999.

[TH95]   J. M. Torrance and L. Hanzo. Comparative study of pilot symbol assisted modem schemes. *Proceedings of Radio Receivers and Associated Systems Conference*, pages 36–41, September 1995.

[TH97]   J.M. Torrance and L. Hanzo. On the upper bound peformance of adaptive QAM in a slow Rayleigh fading. *Electronics Letters*, 32:1218–1230, October 1997.

[TH99]   J. M. Torrance and L. Hanzo. Adaptive modems over slow indoor Rayleigh fading channels. *IEEE Transactions on Vehicular Technology*, 48(4):1237–1251, July 1999.

[THK99]   J. M. Torrance, L. Hanzo, and T. Keller. Interference aspects of adaptive modems over slow Rayleigh fading channels. *IEEE Transactions on Vehicular Technology*, 48(5):1527–1535, September 1999.

IEEE Transactions on Information Theory, 45(2):761–771, March 1999.

[WC99]   L. Welburn and J.K. Cavers. Pilot signals improve the performance of Reed-Solomon erros and erasures decoder in Rayleigh fading channels. *IEEE Transactions on Communications*, 47(5):1527–1535, May 1999.

[WH00]   C.H. Wong and L. Hanzo. Upper-bound performance of a wide-band adaptive modem. *IEEE Transactions on Communications*, 48(3):367–369, March 2000.

[Wol78]   J. Wolfowitz. *Coding Theorems of Information Theory*. Springer-Verlag, 1978.

[WS95]   W.T. Webb and R. Steele. Variable rate QAM for mobile radio. *IEEE Transactions on Communications*, 43:2223–2230, July 1995.

[WYH00] C.H. Wong, B.L. Yeap, and L. Hanzo. Wideband burst-by-burst adaptive modulation with turbo equalization and iterative channel estimation. *Vehicular Technology Conference Proceedings*, 3:2044–2048, Spring 2000.