

Training Optimization for Gauss-Markov Rayleigh Fading Channels

Sami Akin Mustafa Cenk Gursoy

Department of Electrical Engineering

University of Nebraska-Lincoln

Lincoln, NE 68588

Email: sakin1@bigred.unl.edu, gursoy@unl.edu

Abstract—¹ In this paper, pilot-assisted transmission over Gauss-Markov Rayleigh fading channels is considered. A simple scenario, where a single pilot signal is transmitted every T symbols and $T - 1$ data symbols are transmitted in between the pilots, is studied. First, it is assumed that binary phase-shift keying (BPSK) modulation is employed at the transmitter. With this assumption, the training period, and data and training power allocation are jointly optimized by maximizing an achievable rate expression. Achievable rates and energy-per-bit requirements are computed using the optimal training parameters. Secondly, a capacity lower bound is obtained by considering the error in the estimate as another source of additive Gaussian noise, and the training parameters are optimized by maximizing this lower bound.

I. INTRODUCTION

One of the key characteristics of wireless communications that most greatly impact system design and performance is the time-varying nature of the channel conditions, experienced due to mobility and changing physical environment. This has led mainly to three lines of work in the performance analysis of wireless systems. A considerable amount of effort has been expended in the study of cases in which the perfect channel state information (CSI) is assumed to be available at either the receiver or the transmitter or both. With the perfect CSI available at the receiver, the authors in [14] and [15] studied the capacity of fading channels. The capacity of fading channels is also studied in [16] and [17] with perfect CSI at both the receiver and the transmitter. A second line of work has considered fast fading conditions, and assumed that neither the receiver nor the transmitter is aware of the channel conditions (see e.g., [5], [7], [8]). On the other hand, most practical wireless systems attempt to learn the channel conditions but can only do so imperfectly. Hence, it is of great interest to study the performance when only imperfect CSI is available at the transmitter or the receiver. When the channel is not known a priori, one technique that provides imperfect receiver CSI is to employ pilot signals in the transmission to estimate the channel.

Pilot-Assisted Transmission (PAT) multiplexes known training signals with the data signals. These transmission strategies and pilot symbols known at the receiver can be used for channel estimation, receiver adaptation, and optimal decoding [1]. One of the early studies has been conducted by Cavers

in [12], [13] where an analytical approach to the design of PATs is presented. [6] has shown that the data rates are maximized by periodically embedding pilot symbols into the data stream. The amount, placement, and fraction of pilot symbols in the data stream have considerable impact on the data rate. The more pilot symbols are transmitted and the more power is allocated to the pilot symbols, the better estimation quality we have, but the more time for transmission of data is missed and the less power we have for data symbols. Hassibi and Hochwald [10] has optimized the power and duration of training signals by maximizing a capacity lower bound in multiple-antenna Rayleigh block fading channels. An overview of pilot-assisted wireless transmission techniques is presented in [1].

In [2], considering adaptive coding of data symbols without requiring feedback to the transmitter, Abou-Faycal *et al.* studied the data rates achieved with pilot symbol assisted modulation (PSAM) over Gauss-Markov Rayleigh fading channels. In this paper, the training period is optimized by maximizing the achievable rates. The authors in [4] also considered pilot symbol-assisted transmission over Gauss-Markov Rayleigh channels and analyzed the optimal power allocation among data symbols while the pilot symbol has fixed power. They have shown that the power distribution has a decreasing character with respect to the distance to the last sent pilot, and that data power adaptation improves the rates. The authors in [9] considered a similar setting and analyzed training power adaptation but assumed that the power is uniformly distributed among data symbols.

In this paper, considering that no prior channel knowledge is available at the transmitter and the receiver, we focus on a time-varying Rayleigh fading channel. The channel is modeled by a Gauss-Markov model. Pilot symbols which are known by both the transmitter and the receiver are transmitted with a period of T symbols. In this setting, we seek to jointly optimize the training period, training power, and data power allocation by maximizing achievable rates.

II. CHANNEL MODEL

We consider the following model in which a transmitter and a receiver are connected by a time-varying Rayleigh fading channel,

$$y_k = h_k x_k + n_k \quad k = 1, 2, 3, \dots \quad (1)$$

¹This work was supported in part by the NSF CAREER Grant CCF-0546384.

where y_k is the complex channel output, x_k is the complex channel input, h_k and n_k are the fading coefficient and additive noise component, respectively. We assume that h_k and n_k are independent zero mean circular complex Gaussian random variables with variances σ_h^2 and σ_n^2 , respectively. It is further assumed that x_k is independent of h_k and n_k .

While the additive noise samples $\{n_k\}$ are assumed to form an independent and identically distributed (i.i.d.) sequence, the fading process is modeled as a first-order Gauss-Markov process, whose dynamics is described by

$$h_k = \alpha h_{k-1} + z_k \quad 0 \leq \alpha \leq 1, \quad k = 1, 2, 3, \dots, \quad (2)$$

where $\{z_k\}$'s are i.i.d. circular complex Gaussian variables with zero mean and variance equal to $(1-\alpha^2)\sigma_h^2$. In the above formulation, α is a parameter that controls the rate of the channel variations between consecutive transmissions. For instance, if $\alpha = 1$, fading coefficients stay constant over the duration of transmission, whereas, when $\alpha = 0$, fading coefficients are independent for each symbol. For bandwidths in the 10kHz range and Doppler spreads of the order of 100 Hz, typical values for α are between 0.9 and 0.99 [2].

III. PILOT SYMBOL-ASSISTED TRANSMISSION

We consider pilot-assisted transmission where periodically embedded pilot symbols, known by both the sender and the receiver, are used to estimate the fading coefficients of the channel thereby enabling us to track the time-varying channel. We assume the simple scenario where a single pilot symbol is transmitted every T symbols while $T - 1$ data symbols are transmitted in between the pilot symbols. The following average power constraint,

$$\frac{1}{T} \sum_{k=lT}^{(l+1)T-1} E[|x_k|^2] \leq P \quad l = 0, 1, 2, \dots, \quad (3)$$

is imposed on the input. Therefore, the total average power allocated to pilot and data transmission over a duration of T symbols is limited by PT .

Communication takes place in two phases. In the training phase, the pilot signal is sent and the channel output is given by

$$y_{lT} = h_{lT} \sqrt{P_t} + n_{lT} \quad l = 0, 1, 2, 3, \dots \quad (4)$$

where P_t is the power allocated to the pilot symbol. The fading coefficients are estimated via MMSE estimation, which provides the following estimate:

$$\hat{h}_{lT} = \frac{\sqrt{P_t} \sigma_h^2}{P_t \sigma_h^2 + \sigma_n^2} y_{lT}. \quad (5)$$

Following the transmission of the training symbol, data transmission phase starts and $T - 1$ data symbols are sent. Since a single pilot symbol is transmitted, the estimates of the fading coefficients in the data transmission phase are obtained as follows:

$$\hat{h}_k = \frac{\sqrt{P_t} \sigma_h^2}{P_t \sigma_h^2 + \sigma_n^2} \alpha^{k-lT} y_{lT} \quad lT < k \leq (l+1)T - 1. \quad (6)$$

Now, we can express the fading coefficients as

$$h_k = \hat{h}_k + \tilde{h}_k \quad (7)$$

where \tilde{h}_k is the estimation error. Consequently, the input-output relationship in the data transmission phase can be written as

$$y_k = \hat{h}_k x_k + \tilde{h}_k x_k + n_k \quad lT < k \leq (l+1)T - 1. \quad (8)$$

Note that \hat{h}_k and \tilde{h}_k for $lT < k < (l+1)T$ are uncorrelated zero-mean circularly symmetric complex Gaussian random variables with variances

$$\sigma_{\tilde{h}_k}^2 = \frac{P_t \sigma_h^4}{P_t \sigma_h^2 + \sigma_n^2} (\alpha^{k-lT})^2, \quad (9)$$

and

$$\sigma_{h_k}^2 = \sigma_h^2 - \frac{P_t \sigma_h^4}{P_t \sigma_h^2 + \sigma_n^2} (\alpha^{k-lT})^2, \quad (10)$$

respectively.

IV. OPTIMAL POWER DISTRIBUTION AND TRAINING PERIOD FOR BPSK SIGNALS

A. Problem Formulation

In this section, we consider that binary phase-shift keying (BPSK) is employed at the transmitter to send the information. Since our main goal is to optimize the training parameters and identify the optimal power allocation, BPSK signaling is adopted due to its simplicity. In the k^{th} symbol interval, the BPSK signal can be represented by two equiprobable points located at $x_{k,1} = \sqrt{P_{d,k}}$ and $x_{k,2} = -\sqrt{P_{d,k}}$ on the constellation map. Note that $P_{d,k}$ is the average power of the BPSK signal in the k^{th} symbol interval. In this interval, the input-output mutual information conditioned on the value y_{lT} is given by

$$\begin{aligned} I_k(x_k; y_k | y_{lT} = y_{lT}^*) &= \\ &= \frac{1}{2} \int p_{y_k|x_k}(y|x_{k,1}) \log \frac{p_{y_k|x_k}(y|x_{k,1})}{p_{y_k}(y)} dy \\ &+ \frac{1}{2} \int p_{y_k|x_k}(y|x_{k,2}) \log \frac{p_{y_k|x_k}(y|x_{k,2})}{p_{y_k}(y)} dy \end{aligned} \quad (11)$$

where

$$p_{Y_k|X_k}(y_k|x_k) = \frac{1}{\pi(\sigma_{\tilde{h}_k}^2 |x_k|^2 + \sigma_n^2)} \exp\left(\frac{-|y_k - \hat{h}_k x_k|^2}{\sigma_{\tilde{h}_k}^2 |x_k|^2 + \sigma_n^2}\right)$$

and

$$p_{y_k}(y) = \frac{1}{2} p_{y_k|x_k}(y|x_{k,1}) + \frac{1}{2} p_{y_k|x_k}(y|x_{k,2}).$$

We consider the following achievable rate expression, which acts as a lower bound to the channel capacity:

$$I_L(T, P_t, \mathbf{P}_d) = E \left[\frac{1}{T} \sum_{k=lT+1}^{(l+1)T-1} I_k(x_k; y_k | y_{lT} = y_{lT}^*) \right] \quad (12)$$

$$= \frac{1}{T} \sum_{k=lT+1}^{(l+1)T-1} E [I_k(x_k; y_k | y_{lT} = y_{lT}^*)] \quad (13)$$

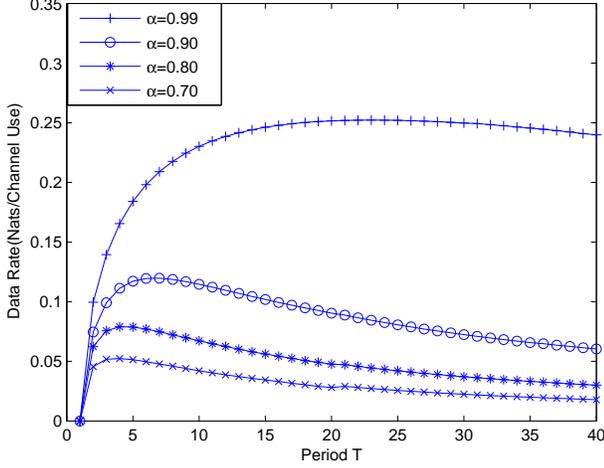


Fig. 1. Achievable data rates vs. training period T for $\alpha = 0.99, 0.90, 0.80$, and 0.70 . $SNR = \frac{P}{\sigma_n^2} = 0dB$

where the expectation is with respect to y_{lT} , and y_{lT}^* is a realization of the random variable y_{lT} . Note that the achievable rate is expressed as a function of the training period, T ; power of the pilot signal, P_t ; and the power allocated to $T - 1$ data symbols transmitted in between the pilot symbols, which is described by the following vector

$$\mathbf{P}_d = [P_{d,1}, P_{d,2}, \dots, P_{d,T-1}]. \quad (14)$$

Our goal is to solve the joint optimization problem

$$(T^*, P_t^*, \mathbf{P}_d^*) = \arg \max_{T, P_t, \mathbf{P}_d} I_L(T, P_t, \mathbf{P}_d) \quad (15)$$

$$P_t + \sum_{k=1}^{T-1} P_{d,k} \leq PT$$

and obtain the optimal training period, and optimal data and pilot power allocations. Since it is unlikely to reach to closed-form solutions, we have employed numerical tools to solve (15).

B. Numerical Results

In this section, we summarize the numerical results. Figure 1 plots the data rates achieved with optimal power allocations as a function of the training period for different values of α . The power level is kept fixed at $P = \sigma_n^2 = 1$. It is observed that the optimal values of the training period, T , are 23, 7, 4, and 4 for $\alpha = 0.99, 0.90, 0.80$, and 0.70 , respectively. Note that the optimal T and optimal data rate are decreasing with the decreasing α . This is expected because the faster the channel changes, the more frequently the pilot symbols should be sent. This consequently reduces the data rates which are already adversely affected by the fast changing and imperfectly known channel conditions. Figures 2 and 3 are the bar graphs providing the optimal training and data power allocation when the training period is at its optimal value. In the graphs, the first bar corresponds to the power of the training symbol while the remaining bars provide the power levels of the data symbols. We immediately observe from both figures that the data symbols, which are farther

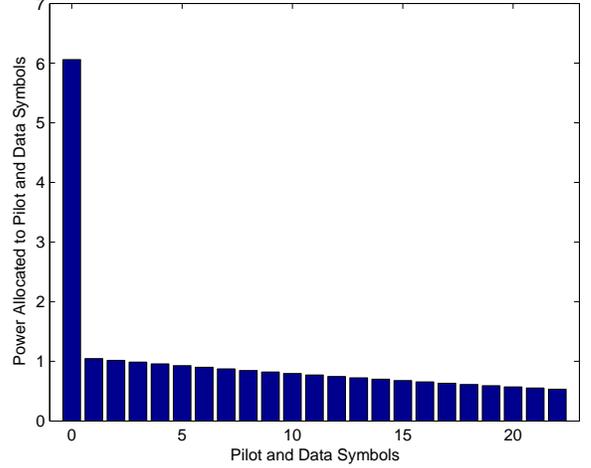


Fig. 2. Optimal power distribution among the pilot and data symbols when $\alpha = 0.99$ and $SNR=0dB$. The optimal period is $T = 23$.

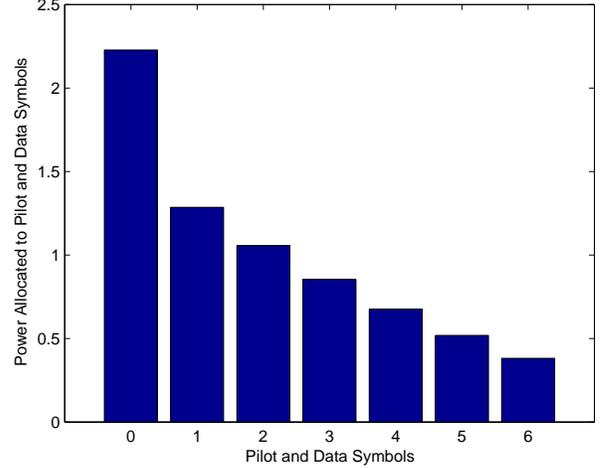


Fig. 3. Optimal power distribution among the pilot and data symbols when $\alpha = 0.90$ and $SNR=0dB$. The optimal period is $T = 7$.

away from the pilot symbol, are allocated less power since channel gets noisier for these symbols due to poorer channel estimates. Moreover, comparing Fig. 2 and Fig. 3, we see that having a longer training period enables us to put more power on the pilot signal and therefore have better channel estimates. We also note that if α is small as in Fig. 3, the power of the data symbols decreases faster as they move away from the pilot symbol. From these numerical results, it is evident that α greatly affects the optimal power allocation and optimal T . Fig. 4 gives the power distribution when $\alpha = 0.90$ and $T = 23$. Note this value of the training period is suboptimal. The inefficiency of this choice is apparent in the graph. Since the channel is changing relatively fast and the quality of the channel estimate deteriorates rather quickly, only half of the available time slots are used for data transmission, leading to a considerable loss in data rates.

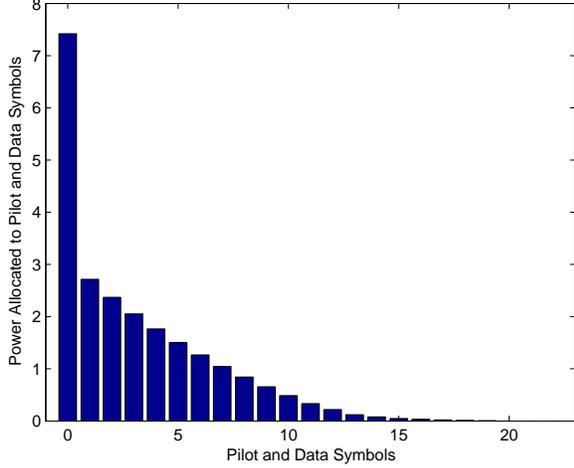


Fig. 4. Optimal power distribution among the pilot and data symbols when $\alpha = 0.90$ and SNR=0dB. The optimal period is $T = 23$.

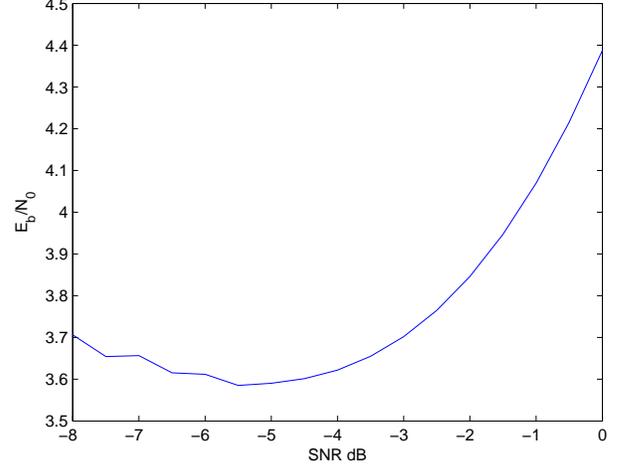


Fig. 5. Bit energy $\frac{E_b}{N_0}$ vs. SNR(dB) when $\alpha = 0.99$.

In systems with scarce energy resources, energy required to send one information bit, rather than data rates, is a suitable metric to measure the performance. The least amount of normalized bit energy required for reliable communications is given by

$$\frac{E_b}{N_0} = \frac{\text{SNR}}{C(\text{SNR})} \quad (16)$$

where $C(\text{SNR})$ is the channel capacity in bits/symbol. In our setting, the bit energy values found from

$$\frac{E_b}{N_0} = \frac{\text{SNR}}{I_L(T^*, P_t^*, \mathbf{P}_d^*)} \quad (17)$$

provide an upper bound on the values obtained from (16), and also gives us indications on the energy efficiency of the system. Fig. 5 plots the required bit energy values as a function of the SNR. The bit energy initially decreases as SNR decreases and achieves its minimum value at approximately SNR= -5.5 dB below which the bit energy requirement starts increasing. Hence, it is extremely energy inefficient to operate below SNR= -5.5 dB. In general, one needs to operate at low SNR levels for improved energy efficiency. From Fig. 6, which plots the optimal training period, T , as a function of the SNR, we observe that T increases as SNR decreases. Hence, training is performed less frequently in the low SNR regime. Fig. 7 provides the pilot and data power allocation when SNR = -7dB, $\alpha = 0.99$, and $T = 65$. It is interesting to note that although T is large, a considerable portion of the available time slots are not being used for transmission. This approach enables the system to put more power on the pilot symbol and nearby data symbols. Hence, although the system trains and transmits less frequently, a more accurate channel estimate is obtained and used in return.

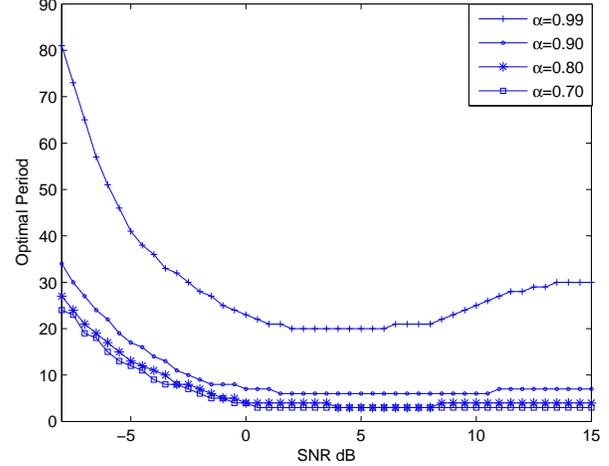


Fig. 6. Optimal training period T vs. SNR for $\alpha = 0.99, 0.90, 0.80, 0.70$.

V. LOW COMPLEXITY TRAINING OPTIMIZATION

Recall that the input-output relationship in the data transmission phase is given by²

$$y_k = \hat{h}_k x_k + \tilde{h}_k x_k + n_k \quad k = 1, 2, \dots, T-1. \quad (18)$$

In the preceding section, we fixed the modulation format and computed the input-output mutual information achieved in the channel (18). In this section, we pursue another approach akin to that in [10]. We treat the error in the channel estimate as another source of additive noise and assume that

$$w_k = \tilde{h}_k x_k + n_k \quad (19)$$

is zero-mean Gaussian noise with variance

$$\sigma_{w_k}^2 = \sigma_{\tilde{h}_k}^2 P_{d,k} + \sigma_n^2. \quad (20)$$

²It is assumed that a single pilot signal is transmitted at $k = 0$.

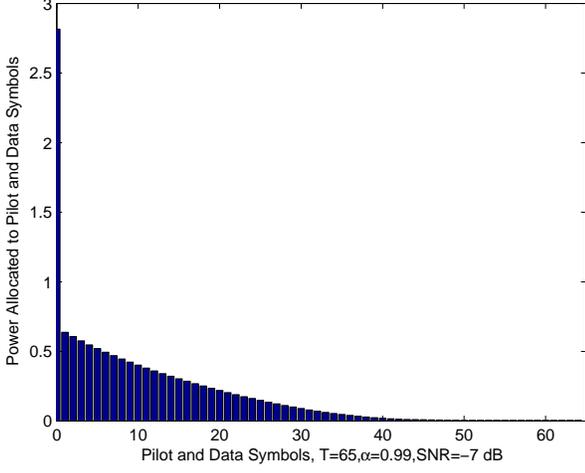


Fig. 7. Optimal power distribution for the pilot and data symbols when $\alpha = 0.99$ and SNR=-7dB. The optimal period is $T=65$.

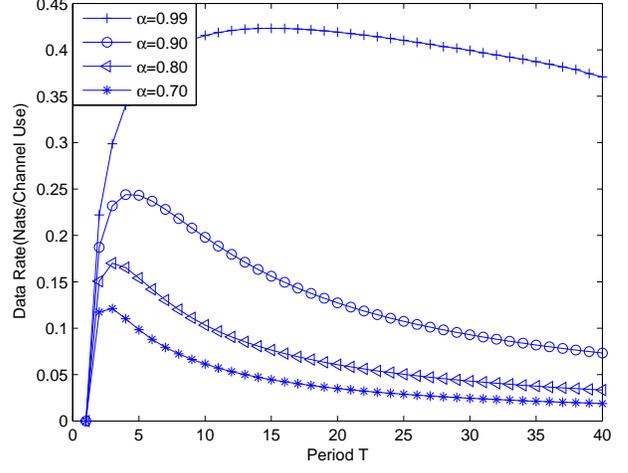


Fig. 8. Achievable data rates vs. training period T for $\alpha = 0.99, 0.90, 0.80,$ and 0.70 . SNR=5dB

where $P_{d,k} = E[|x_k|^2]$ is the average power of the symbol x_k and $\sigma_{h_k}^2$ is given in (10). Since the Gaussian noise is the worst case noise [10], the capacity of the channel

$$y_k = \hat{h}_k x_k + w_k \quad k = 1, 2, \dots \quad (21)$$

is a lower bound to the capacity of the channel given in (18). An achievable rate expression for channel (21) is

$$I_{\text{worst}} = \max_{T, P_t} \max_{\mathbf{x}} \frac{1}{T} \sum_{k=1}^{T-1} I_k(x_k; y_k | \hat{h}_k) \quad (22)$$

$$= \max_{T, P_t} \max_{\substack{\mathbf{P}_d \\ P_{d,k} \geq 0 \forall k \\ \sum_{k=1}^{T-1} P_{d,k} \leq PT - P_t}} \frac{1}{T} \sum_{k=1}^{T-1} \max_{x_k} I_k(x_k; y_k | \hat{h}_k) \quad (23)$$

$$= \max_{T, P_t, \mathbf{P}_d} \frac{1}{T} \sum_{k=1}^{T-1} E \left[\log \left(1 + \frac{\sigma_{h_k}^2 P_{d,k}}{\sigma_{h_k}^2 P_{d,k} + \sigma_n^2} |\xi|^2 \right) \right] \quad (24)$$

In (22), $\mathbf{x} = (x_1, x_2, \dots, x_{T-1})$ denotes the vector of $T-1$ input symbols, and the inner maximization is over the space of joint distribution functions of \mathbf{x} . (23) is obtained by observing that once the data power distribution is fixed, the maximization over the joint distribution can be broken down into separate maximization problems over marginal distributions. (24) follows from the fact that Gaussian input maximizes the mutual information $I(x_k; y_k | \hat{h}_k)$ when the channel in consideration is (21). Note that in (24), ξ is a zero mean, unit variance, circular complex Gaussian random variable, and the expectation is with respect to ξ . We can again numerically solve the above optimization and Fig. 8 plots the achievable data rates with optimal power allocation as a function of T for different values of α when SNR=5dB. An even simpler optimization problem

results if we seek to optimize the upper bound

$$\frac{1}{T} \sum_{k=1}^{T-1} E \left[\log \left(1 + \frac{\sigma_{h_k}^2 P_{d,k}}{\sigma_{h_k}^2 P_{d,k} + \sigma_n^2} |\xi|^2 \right) \right] \quad (25)$$

$$\leq \frac{1}{T} \sum_{k=1}^{T-1} \log \left(1 + \frac{\sigma_{h_k}^2 P_{d,k}}{\sigma_{h_k}^2 P_{d,k} + \sigma_n^2} \right), \quad (26)$$

which is obtained by using the Jensen's inequality and noting that $E[|\xi|^2] = 1$. In this case, the optimization problem becomes

$$\max_{T, P_t, \mathbf{P}_d} \frac{1}{T} \sum_{k=1}^{T-1} \log \left(1 + \frac{\sigma_{h_k}^2 P_{d,k}}{\sigma_{h_k}^2 P_{d,k} + \sigma_n^2} \right) \quad (27)$$

$$= \max_{T, P_t, \mathbf{P}_d} \frac{1}{T} \log \left(\prod_{k=1}^{T-1} \left(1 + \frac{\sigma_{h_k}^2 P_{d,k}}{\sigma_{h_k}^2 P_{d,k} + \sigma_n^2} \right) \right). \quad (28)$$

Since logarithm is a monotonically increasing function, the optimal training and data power allocation for fixed T can be found by solving

$$\max_{T, P_t, \mathbf{P}_d} \prod_{k=1}^{T-1} \left(1 + \frac{\sigma_{h_k}^2 P_{d,k}}{\sigma_{h_k}^2 P_{d,k} + \sigma_n^2} \right). \quad (29)$$

It is very interesting to note that the optimal power distribution found by solving (29) is very similar to that obtained from (15) where BPSK signals are considered. Figure 9 plots the achievable data rates as a function of training period when BPSK signals are employed for transmission. Hence, the data rates are computed using (12). In the figure, the solid line shows the data rates achieved with power distribution found from (15) while the dashed line corresponds to rates achieved with power allocation obtained from (29). Note that both curves are very close and the training period is maximized at approximately the same value.

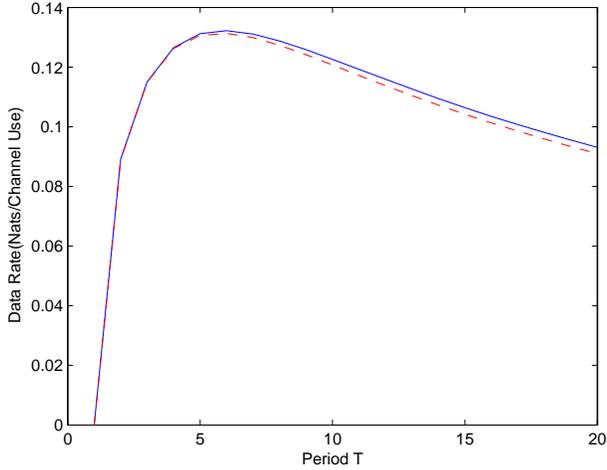


Fig. 9. Achievable data rates for BPSK signals vs. training period T for $\alpha = 0.90$. SNR=0dB.

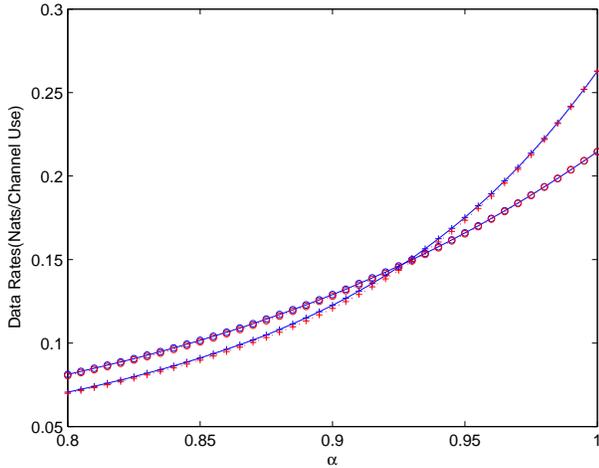


Fig. 10. Achievable data rates for BPSK signals vs. α for $T = 6$ and 10 . SNR = 0 dB. "+" and solid line" and "+" and dotted line" are plotting rates achieved with power allocation from (29) and (15), respectively, when $T = 10$. "o and solid line" and "o and dotted line" are plotting rates achieved with power allocation from (29) and (15), respectively, when $T = 6$.

Fig. 10 plots the achievable rates for BPSK signals as a function of the parameter α for $T = 6$ and 10 . The power distribution is again obtained from both (29) and (15). We again recognize that the loss in data rates is negligible when (29) is used to find the power allocation.

VI. CONCLUSION

We have studied the problem of training optimization in pilot-assisted wireless transmissions over Gauss-Markov Rayleigh fading channels. We have considered a simple scenario where a single pilot is transmitted every T symbols for channel estimation and $T - 1$ data symbols are transmitted in between the pilot symbols. MMSE estimation is employed to estimate the channel. We have jointly optimized the train-

ing period, T , and data and training power distributions by maximizing achievable rate expressions. We have provided numerical results showing the optimal parameters, power distributions, and maximized achievable rates. We have also studied the energy efficiency of pilot-assisted transmissions by considering the energy-per-bit requirements.

REFERENCES

- [1] L. Tong, B.M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions," *IEEE Signal Processing Magazine*, vol. 21, no. 06, pp. 12-25, November 2004.
- [2] I. Abou-Faycal, M. Médard, and U. Madhow, "Binary adaptive coded pilot symbol assisted modulation over Rayleigh fading channels without feedback," *IEEE Trans. Commun.*, vol. 53, pp. 1036-1046, June 2005.
- [3] M. Médard, I. Abou-Faycal, and U. Madhow, "Adaptive coded modulation without channel feedback for pilot symbol assisted modulation," *38th Annual Allerton Conference on Communication, Control and Computing, 2000*
- [4] A. Bdeir, I. Abou-Faycal, and M. Médard, "Power allocation schemes for pilot symbol assisted modulation over rayleigh fading channels with no feedback," *Communications, IEEE International Conference on*, vol. 2, pp. 737-741, June 2004.
- [5] I. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete-time memoryless Rayleigh fading channels," *IEEE Trans. Inform. Theory*, Vol. 47, pp.1290-1301, May 2001
- [6] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for unknown channels," in *Conference on Information Sciences and Systems*, March 21-23, 2001.
- [7] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139-157, Jan. 1999.
- [8] M. C. Gursoy, H. V. Poor, and S. Verdú, "The noncoherent Rician fading channel - Part I : Structure of the capacity-achieving input," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2193-2206.
- [9] M. F. Sencan, and M. C. Gursoy, "Achievable rates for pilot-assisted transmission over Rayleigh fading channels," *Proceedings of the 40th Annual Conference on Information Science and Systems, Princeton University, Princeton, NJ, March 22-24 2006*
- [10] B. Hassibi, and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inform. Theory*, Vol. 49, pp.951-963, April 2003
- [11] S. Ohno, and G.B. Giannakis, "average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 712-720, October 2002.
- [12] J.K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Vehicular Tech.*, vol. 40, pp. 686-693, November 1991.
- [13] J.K. Cavers, "Pilot assisted symbol modulation and differential detection in fading and delay spread," *IEEE Trans. Inform. Theory*, vol. 43, no. 7, pp. 2206-2212, 1995.
- [14] T. Ericson, "A gaussian channel with slow fading," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 353-355, May 1970.
- [15] W.C.Y. Lee, "Estimate of channel capacity in rayleigh fading environment," *IEEE Trans. Vehic. Technol.*, vol. 39, no. 3, pp. 187-189, August 1990.
- [16] A.J. Goldsmith and P.P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, no. 06, pp. 1986-1992, November 1997.
- [17] A.J. Goldsmith and M.S. Alouini, "Comparison of fading channel capacity under different csi assumptions," *Proc. IEEE Vehic. Technol. Conf.*, pp. 1844-1848, September 2000.