How Much Training is Needed in Fading Multiple Access Channels?

Jonathan Scarlett, Jamie Evans and Subhrakanti Dey
Department of Electrical and Electronic Engineering
University of Melbourne, Victoria, Australia

***Abstract*—We optimize the tradeoff between multiuser diversity and training overhead in a single antenna narrowband multiple access channel with a large number of users. A block fading model with independent Rayleigh gains is considered, in which training sequences are sent to the base station one at a time and the user estimated to be the strongest is scheduled to transmit. Considering a lower bound on the ergodic sum capacity with channel uncertainty under an average total power constraint, we optimize the proportion of time and power spent on training in each block. By analyzing the asymptotic behavior of the system as the block length grows large, we optimize the number of users considered for transmission in each block with respect to an approximate expression for the achievable rate, and find first order expressions for the resulting parameters.**

I. Introduction

Multiuser diversity is a well known technique for taking advantage of channel fluctuations in wireless communication systems [1], [2]. In a cell with a large number of users experiencing independent fading, high rates of communication can be obtained by scheduling only the users with the strongest channels. More specifically, in a multiple access channel (MAC) with an average total power constraint, symmetric fading statistics and full channel state information (CSI), ergodic sum capacity is maximized by allowing only the strongest user to transmit, with the power allocation given by waterfilling [1]. Furthermore, when the tail of the fading distribution satisfies certain conditions, the ergodic sum capacity scales as $\log\log K_{\text{total}}$, where K_{total} is the total number of users in the system [2]. In particular, this result holds for channel distributions with exponential tails, such as the Rayleigh distribution.

In practical systems, full CSI is an unreasonable assumption, and channel estimates are instead obtained via training. This can require significant overhead in terms of both time and power, particularly when the number of users in the system is large. In time division duplex (TDD) systems, this overhead can be greatly reduced by taking advantage of channel reciprocity [2]. However, we consider a frequency division duplex (FDD) system in which such techniques cannot be used. Given a finite coherence time, there is a limit to how long is spent on training before the channel estimates become stale, and hence a limit on how many users can train the base station during this time. Consequently, the ergodic sum capacity remains bounded as the total number of users in the system grows large, and $\log\log K_{\text{total}}$ scaling of capacity is not achieved.

In this paper, we consider a narrowband single antenna MAC with block fading and independent Rayleigh distributed channel coefficients. During each block, K users train the base station one at a time, after which the base station selects the user with the strongest channel and feeds back the corresponding index. We aim to maximize a lower bound on the ergodic capacity with respect to the training time, training power and number of users considered for transmission, considering losses due to reduced degrees of freedom, reduced power, and channel uncertainty. An extended version of this paper is given in [3], where we provide a more detailed analysis and further discussions.

Our approach is similar to [4], where training time and power are optimized along with the number of subchannels trained in a single-user wideband system. In [5], the work of [4] is extended to the multiuser wideband case with random training sequences, under the assumption that the number of users grows linearly with the block length. That is, optimization is done over the number of subchannels for a fixed number of users but not vice versa. Analysis of a multiuser narrowband system is performed in [6], but with a focus on the downlink channel, assuming that users estimate their own channel perfectly, with perfect feedback to the base station requiring a fixed number of bits per user. Much of our notation is borrowed from [4].

II. System Model and Problem Statement

We consider a single antenna FDD narrowband MAC with K_{total} users communicating with a base station, where K_{total} is very large. The transmitted data is assumed to be delay-insensitive. The channel is modeled as a Rayleigh block fading channel with L symbols per block and independent fades between blocks. Within each block, K users are considered for transmission. We assume that K_{total} is always greater than K, with the group of users considered varying between blocks for fairness (e.g. using round robin selection). Under this setup, the system is described by

$$\mathbf{y} = \sum_{k=1}^{K} h_k \mathbf{x}_k + \mathbf{z}$$

where $\mathbf{y}$ is the $L \times 1$ received signal vector, $\mathbf{x}_k$ is the $L \times 1$ transmit symbol vector for user k, $h_k \overset{d}{=} CN(0, \sigma_h^2)$ is the channel coefficient of user k, and $\mathbf{z} \overset{d}{=} CN(\mathbf{0}, \sigma_z^2 \mathbf{I})$ is an $L \times 1$ vector of noise samples (here $\overset{d}{=}$ means "distributed as" and $CN(\cdot,\cdot)$ is the complex Gaussian distribution). The

978-1-61284-402-2/11/$26.00 ©2011 IEEE

transmitted symbols are subject to an average total power constraint, $E[\frac{1}{L}\sum_{k=1}^{K}||\mathbf{x}_k||^2] \leq P$. The users are assumed to be synchronized with their coherence blocks aligned in time, and each user is assumed to experience independent fading. We note that due to the symmetry of the setup, the power constraint could be replaced by a more realistic *individual* average power constraint of $\frac{P}{K_{\text{total}}}$ for each of the K_{total} users without affecting the analysis. However, the analysis of an asymmetric setup with individual power constraints is beyond the scope of this paper.

Since the channel coefficients h_k are unknown at the base station, the first T symbols of each coherence block are used for training. One at a time, the K users under consideration transmit a training sequence containing $\overline{T}$ symbols, giving a total of $T = K\overline{T}$ symbols dedicated to training. Each user transmits with power P_T when sending their own training sequence, and remains silent while the other training sequences are sent. At the base station, a minimum mean square error (MMSE) channel estimate $\widehat{h}_k$ is obtained for each user, with the corresponding channel estimation error denoted by $e_k = h_k - \widehat{h}_k$. The variance of this error is given by [4]

$$\sigma_e^2 = E[|e_k|^2] = \sigma_h^2 - \sigma_{\widehat{h}}^2 = \sigma_h^2\left(1 - \frac{\sigma_h^2\overline{T}P_T}{\sigma_h^2\overline{T}P_T + \sigma_z^2}\right) \quad (1)$$

where $\sigma_{\widehat{h}}^2 = E[|\widehat{h}_k|^2]$ is the variance of $\widehat{h}_k$, which is equal for each of the trained users.

Since the system is narrowband and single-antenna, it is preferable to schedule only one user in each block [1], [3]. Hence, the base station schedules the user with the strongest channel estimate, $\max_{k=1,\dots,K}|\widehat{h}_k|^2$, which will be denoted as $|\widehat{h}^*|^2$. We assume for simplicity that the feedback is instantaneous and error-free, hence the selected user uses the remaining $L - T$ symbols of the block for data transmission. The average transmit power during this time is fixed at P_D, which is obtained from

$$P_D = \frac{P - \alpha P_T}{1 - \alpha} \quad (2)$$

where $\alpha = \frac{T}{L}$ is the fraction of the coherence time dedicated to training. That is, P_D is chosen so that the average total power constraint is met with equality. While a fixed data transmit power is generally suboptimal, it achieves performance very close to optimal waterfilling even for moderate values of K [7], while being simple to analyze and having a low feedback requirement.

Since the ergodic sum capacity of a fading channel with uncertainty is not yet known, we instead use the lower bound

$$\underline{C} = (1-\alpha)E\left[\log\left(1 + \frac{P_D|\widehat{h}^*|^2}{P_D\sigma_e^2 + \sigma_z^2}\right)\right] \quad (3)$$

which is achieved by treating the channel estimation error as additive Gaussian noise [8]. The problem is to optimize the fraction of time spent training α, training power P_T, and number of users K in order to maximize $\underline{C}$. The optimal parameters will be denoted by α^*, P_T^* and K^*, and the corresponding achievable rate by $\underline{C}^*$.[1] While optimizing a lower bound on capacity may not give exactly the same results as optimizing the true capacity, this problem still provides valuable insight into the tradeoff between multiuser diversity and training overhead. Spending more time and power on training will clearly reduce the estimation error, but at the expense of reducing the time and power left for data transmission. Similarly, considering more users in each coherence block will increase the capacity via multiuser diversity, but at the expense of the requirement of additional training.

III. Optimization

We write the achievable rate in two equivalent forms,

$$\underline{C} = (1-\alpha)\int_0^\infty \log\left(1 + \frac{(P-\epsilon_T)t}{(P-\epsilon_T)\sigma_e^2 + \sigma_z^2(1-\alpha)}\right) f(t)dt \quad (4)$$

$$\underline{C} = (1-\alpha)E\left[\log\left(1 + \frac{1}{x}|h_1^*|^2\right)\right] \quad (5)$$

where $f(t)$ is the cumulative distribution function of $|\widehat{h}^*|^2$, $\epsilon_T = \alpha P_T$, $|h_1^*|^2$ is the maximum of K independent $\exp(1)$ random variables, and

$$x = \frac{P_D\sigma_e^2 + \sigma_z^2}{P_D\sigma_{\widehat{h}}^2} \quad (6)$$

is the *effective inverse signal to noise ratio*. In order to obtain expressions for the optimal proportion of time and power spent on training for a given K, we apply the techniques of [4] to the multiuser setting. We begin by optimizing α for fixed values of ϵ_T and K.[2] From (1), and writing $\overline{T}P_T = \frac{L}{K}\epsilon_T$, σ_e^2 and $\sigma_{\widehat{h}}^2$ depend on α only through ϵ_T. Hence, from (4), optimizing α is equivalent to maximizing $(1-\alpha)\log(1 + \frac{a}{b-\alpha})$ for some $a, b > 0$. This function is decreasing in α, hence we choose α to be as low as possible while still ensuring all K users perform training. This is achieved by

$$\alpha^* = \frac{K}{L}. \quad (7)$$

by setting $\overline{T} = 1$ training symbol per user. That is, for *any* value of ϵ_T it is optimal to choose $\overline{T} = 1$.[3]

Next we optimize the training power. Instead of optimizing P_T directly, we optimize the *proportion* of power spent on training, denoted by $\bar{\epsilon}_T$ and given by $\bar{\epsilon}_T = \frac{\epsilon_T}{P}$. From (5) it is clear that $\underline{C}$ is decreasing in x for any fixed K. Hence the optimal value of $\bar{\epsilon}_T$, denoted by $\bar{\epsilon}_T^*$, minimizes x. Substituting (1) and (2) into (6) and setting $\overline{T} = 1$ gives

$$x = \left(1 + \frac{\alpha}{S\bar{\epsilon}_T}\right)\left(1 + \frac{(1-\alpha)}{S(1-\bar{\epsilon}_T)}\right) - 1 \quad (8)$$

[1] In general these will implicitly be a function of K (e.g. $\alpha^* = \frac{K}{L}$ in (7)).

[2] While ϵ_T depends on α, it can be kept fixed as α varies by adjusting P_T accordingly. This corresponds to keeping the training *energy* fixed while varying the training time and power.

[3] Identical performance could also be obtained using any orthogonal training sequences of length K (e.g. Walsh-Hadamard sequences) with MMSE estimation

where $S = \frac{P\sigma_h^2}{\sigma_z^2}$ is the overall signal to noise ratio (SNR). Setting $\frac{\delta x}{\delta \bar{\epsilon}_T} = 0$ gives $\bar{\epsilon}_T^*$ as the solution of a quadratic equation, the positive solution of which is $\bar{\epsilon}_T^* = \frac{1}{2}$ when $\alpha = \frac{1}{2}$, and

$$\bar{\epsilon}_T^* = \frac{1}{S(1-2\alpha)}\big(-(\alpha(S+1)-\alpha^2)+ \sqrt{\alpha(S+S^2)+(1-S-S^2)\alpha^2-2\alpha^3+\alpha^4}\big) \quad (9)$$

when $\alpha \neq \frac{1}{2}$. With α^* and $\bar{\epsilon}_T^*$ known in closed form for any given K, K^* can be found using an exhaustive search over $K \in \{1, 2, ..., L-1\}$.

IV. Scaling

In this section we present the asymptotic behavior of α^*, $\bar{\epsilon}_T^*$ and P_T^* as $L \to \infty$, and approximate K^* by optimizing over a suitable approximation of $\underline{C}$. For two functions $f(L)$ and $g(L)$, we write $f = O(g)$ if $|f| \leq c|g|$ for some constant c when L is sufficiently large, $f = o(g)$ if $\lim_{L\to\infty} \frac{f}{g} = 0$, $f = \Theta(g)$ if $f = O(g)$ and $f \neq o(g)$, and $f \sim g$ if $\lim_{L\to\infty} \frac{f}{g} = 1$. Due to space constraints, we give only outlines of some of the proofs, referring the reader to [3] for details.

We begin with a lemma showing that the achievable rate is unbounded for large L, and that K^* grows sublinearly with L.

Lemma 1. *As the block length L grows large, $\underline{C}^* \to \infty$ and $\alpha^* \to 0$.*

Proof: Suppose that the chosen parameters are $K = L^{1/2}$ and $\bar{\epsilon}_T = L^{-1/4}$. Using $\alpha = \frac{K}{L}$ we have $\alpha \to 0$, and from (8) we obtain $x \sim \frac{1}{S}$. Substituting these into (5) gives $\underline{C} \sim E[\log(1+S|h_1^*|^2)]$. The right hand side of this asymptotic expression corresponds to the ergodic capacity of a MAC with Rayleigh fading, K users and zero estimation error, which implies $\underline{C} \sim \log\log K$. Substituting $K = L^{1/2}$ gives $\underline{C} \sim \log\log L$, which proves that $\underline{C} \to \infty$ is achievable and therefore $\underline{C}^* \to \infty$.

To prove that $\alpha^* \to 0$, we note that even if perfect channel estimation is assumed with the only effect of training being a loss in temporal degrees of freedom, the achievable rate scales as $(1-\alpha)\log\log K \leq (1-\alpha)\log\log L$, where the inequality follows from $K \leq L$. Since $\underline{C}$ is a lower bound on this rate it is clear that $\alpha \neq o(1)$ is suboptimal, since we have shown that $\underline{C} \sim \log\log L$ is achievable. ■

Since $\alpha^* \to 0$ by Lemma 1, meaningful expressions for the parameters are obtained by considering only the lowest powers of $\alpha^* = \frac{K}{L}$, or the highest powers of $\frac{L}{K}$. Using this result, we give second order asymptotic expressions for $\bar{\epsilon}_T^*$ and P_T^* in terms of K and L.

Lemma 2. *As $K \to \infty$ and $L \to \infty$ with $K = o(L)$, $\bar{\epsilon}_T^*$ and P_T^* satisfy*

$$\bar{\epsilon}_T^* \sim \sqrt{\frac{S+1}{S}}\sqrt{\frac{K}{L}} \quad (10)$$

$$P_T^* \sim P\sqrt{\frac{S+1}{S}}\sqrt{\frac{L}{K}} \quad (11)$$

and the corresponding estimation error satisfies

$$(\sigma_e^*)^2 \sim \frac{\sigma_z^2}{P}\sqrt{\frac{S}{S+1}}\sqrt{\frac{K}{L}}. \quad (12)$$

Proof: Keeping only the most significant terms of (9) and substituting $\alpha^* = \frac{K}{L}$, (10) follows. Substituting (10) into $P_T^* = \frac{\bar{\epsilon}_T^* P}{\alpha^*} = \frac{\bar{\epsilon}_T P L}{K}$ gives (11). Finally, simplifying (1) as $\sigma_e^2 \sim \frac{\sigma_z^2}{P_T}$ and substituting (11) gives (12). ■

In order to obtain expressions for each of the parameters in terms of L alone, optimization over K is required. However, $\underline{C}$ appears to be difficult to optimize over K directly. To simplify the analysis, we consider two approximations of $\underline{C}$, given by

$$\underline{C}_{a1} = \big(1 - \frac{K}{L}\big)\log\left(1 + \frac{1}{x}\log K\right) \quad (13)$$

$$\underline{C}_{a2} = \big(1 - \frac{K}{L}\big)\log\left(1 + S\big(1 - 2\sqrt{\frac{S+1}{S}}\sqrt{\frac{K}{L}}\big)\log K\right). \quad (14)$$

We denote the value of K which maximizes $\underline{C}_{a2}$ as K_a^*. While we do not claim that K_a^* and K^* have the exact same behavior, the following lemma shows that asymptotically there is zero loss in the rate achieved by optimizing $\underline{C}_{a1}$ or $\underline{C}_{a2}$ instead of $\underline{C}$. The accuracy of these approximations is further verified via numerical results in Section V.

Lemma 3. *Suppose α and $\bar{\epsilon}_T$ are chosen according to* (7) *and* (9) *respectively. If K is chosen to optimize any one of $\underline{C}$, $\underline{C}_{a1}$ or $\underline{C}_{a2}$ then $\lim_{L\to\infty} |\underline{C} - \underline{C}_{a1}| = 0$ and $\lim_{L\to\infty} |\underline{C} - \underline{C}_{a2}| = 0$.*

Proof: The following upper and lower bounds can be obtained using Jensen's inequality and Markov's inequality respectively (see Appendix A of [3] for details)

$$\underline{C} \leq \underline{C}_{a1} + (1-\alpha)\log\left(1 + O\big(\frac{1}{x + \log K}\big)\right) \quad (15)$$

$$\underline{C} \geq \underline{C}_{a1} + (1-\alpha)\log\left(1 + o\big(\frac{\log K}{x + \log K}\big)\right) + O\left(\frac{\log\big(1 + \frac{1}{x}\log K\big)}{\log K}\right). \quad (16)$$

Substituting $\alpha = \alpha^*$ and $\bar{\epsilon}_T = \bar{\epsilon}_T^*$ into (8) gives $x \sim \frac{1}{S}$. Hence, applying $x = O(1)$ and $K \to \infty$, both of these bounds simplify to $\underline{C}_{a1} + o(1)$. Combining these, it follows that $\lim_{L\to\infty} |\underline{C} - \underline{C}_{a1}| = 0$.

The expression for $\underline{C}_{a2}$ is obtained by substituting the asymptotic expressions for α^* and $\bar{\epsilon}_T^*$ into $\underline{C}_{a1}$ and performing asymptotic simplifications. In [3] we show that these simplifications not only lead to zero asymptotic loss in capacity, but also have no effect on the first and second order expressions for K_a^*. ■

Using $\underline{C}_{a2}$, we can now find an expression for L in terms of K_a^*.

Lemma 4. *K_a^* satisfies*

$$L \sim \frac{S+1}{S} K_a^* (\log K_a^*)^2. \tag{17}$$

Proof: Using (14) and setting $\frac{\delta}{\delta K}\underline{C}_{a2} = 0$ gives the necessary condition for K to maximize $\underline{C}_{a2}$,

$$\frac{S(L-K)\left(2\big(1-c\sqrt{\frac{K}{L}}\big) - c\sqrt{\frac{K}{L}}\log K\right)}{2K\left(S\big(1-c\sqrt{\frac{K}{L}}\big)\log K + 1\right)} = \log\left(1 + S\big(1-c\sqrt{\frac{K}{L}}\big)\log K\right) \tag{18}$$

where $c = 2\sqrt{\frac{S+1}{S}}$. Hence,

$$\frac{L\big(2 - c\sqrt{\frac{K}{L}}\log K\big)}{2K\log K} \sim \log\log K. \tag{19}$$

Defining

$$\rho = \sqrt{\frac{K}{L}}\log K \tag{20}$$

we obtain

$$L = \frac{1}{\rho^2} K (\log K)^2. \tag{21}$$

Substituting (20) and (21) into (19) gives $\frac{1}{\rho^2}(1-\frac{\rho c}{2})\log K \sim \log\log K$, which is only possible if $\rho \sim \frac{2}{c}$. Therefore, $L \sim \frac{c^2}{4}K(\log K)^2$. Substituting $c = 2\sqrt{\frac{S+1}{S}}$ concludes the proof. ■

Combining Lemmas 2 and 4, the following theorem gives first order expressions for the optimized parameters in terms of L, as well as the corresponding estimation error and achievable rate.

Theorem 5. *The first order asymptotic expression for K_a^* is*

$$K_a^* \sim \frac{S}{S+1}\frac{L}{(\log L)^2} \tag{22}$$

and with $K = K_a^*$ the optimal parameters satisfy

$$\alpha^* \sim \frac{S}{S+1}\frac{1}{(\log L)^2} \tag{23}$$

$$\bar{\epsilon}_T^* \sim \frac{1}{\log L} \tag{24}$$

$$P_T^* \sim \frac{P(S+1)}{S}\log L \tag{25}$$

with corresponding estimation error and achievable rate, respectively, given by

$$(\sigma_e^*)^2 \sim \frac{\sigma_z^2}{P}\frac{1}{\log L} \tag{26}$$

$$\underline{C}^* \sim \log\log L. \tag{27}$$

Proof: From (17) we obtain $\log L \sim \log K_a^*$. Combining this with (17), the expression for K_a^* follows. The expression for α^* follows immediately from $\alpha^* = \frac{K}{L}$. Taking the square root gives $\sqrt{\frac{K}{L}} \sim \sqrt{\frac{S}{S+1}}\frac{1}{\log L}$, which can be substituted into (10), (11), (12) and (14) to obtain the expressions for $\bar{\epsilon}_T^*$, P_T^*, $(\sigma_e^*)^2$ and $\underline{C}^*$ respectively. ■

From these results, we see that the proportion of time and power spent on training tend to zero at rates $O(\frac{1}{(\log L)^2})$ and $O(\frac{1}{\log L})$ respectively. The transmit power during training increases as $O(\log L)$, and the estimation error decreases as $O(\frac{1}{\log L})$. The number of users increases as $O(\frac{L}{(\log L)^2})$, and the achievable rate as $O(\log\log L)$. Hence, the amount of multiuser diversity achieved depends primarily on the block length, rather than the total number of users in the system.

V. Numerical Results

In this section we present numerical results of the system. We use $P = 1$, $\sigma_h^2 = 1$ and $\sigma_z^2 = 0.1$, giving an overall SNR of $S = 10$. Figure 1 shows the values of $\underline{C}$ versus K with the block length fixed at $L = 250$. Even with this relatively small block length, only a small proportion of the time is spent training, with the optimal number of users at $K^* = 15$.

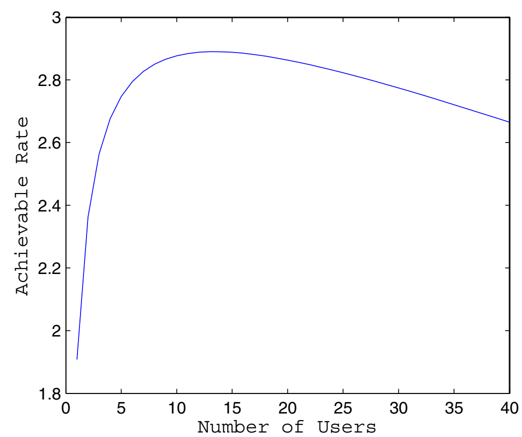

Figure 1. Achievable rate as a function of K with $L = 250$

Plots of α^*, P_T^* and K^* are shown in Figures 2, 3 and 4 respectively. The first order order asymptotic expressions derived in Section IV are shown on the same axes, and the second order expressions from [3] are shown for completeness. We see that the first order expressions have the same asymptotic growth rate as the true values, and the second order expressions follow the true optimal values much more closely. This suggests that $\underline{C}_{a2}$ in (14) is a valid approximation for $\underline{C}$, and that the asymptotic expressions obtained provide good insight into the optimal parameters even for moderate values of L.

VI. Conclusion

We have analyzed a single antenna FDD narrowband MAC with training and best-user feedback. Using a Rayleigh block fading channel model with independent fading between users, a closed form expression has been computed for the optimal proportion of power spent on training, and it has been shown that the optimal training sequence length is $\overline{T} = 1$ symbol

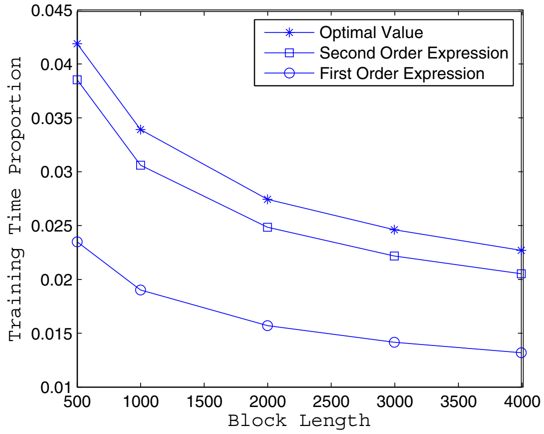

Figure 2. Optimal values and asymptotic expressions for P_T

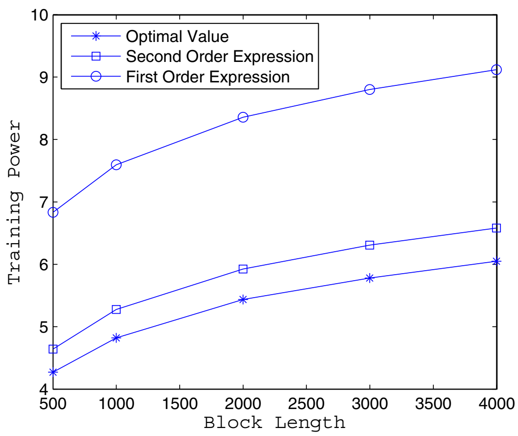

Figure 3. Optimal values and asymptotic expressions for P_T

per user. The asymptotic behavior of the parameters at large block lengths has been analyzed. By optimizing with respect an approximate expression for the achievable rate, first order asymptotic expressions have been obtained for the number of users considered in each block, the proportion of time and power spent training, the average training power, and the corresponding estimation error and achievable rate. Possible further work includes the multiple-antenna setting and fading distributions other than Rayleigh.

REFERENCES

[1] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *IEEE International Conference on Communications*, Seattle, WA, June 1995.

[2] X. Qin and R. A. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *IEEE INFOCOM*, San Francisco, CA, March 2003.

[3] J. Scarlett, J. Evans, and S. Dey, "On the tradeoff between multiuser diversity and training overhead in multiple access channels," arXiv:1105.3531v2 [cs.IT].

[4] M. Agarwal and M. L. Honig, "Wideband fading channel capacity with training and partial feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4865–4873, October 2010.

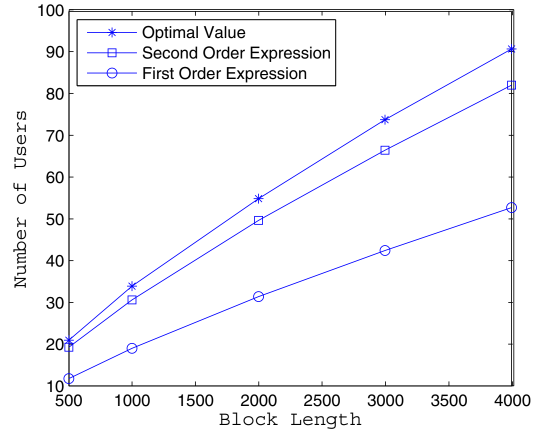

Figure 4. Optimal values and asymptotic expressions for K

[5] ——, "Spectrum sharing on a wideband fading channel with limited feedback," in *CrownCom International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Orlando, FL, August 2007.

[6] A. Rajanna and N. Jindal, "Multiuser diversity in downlink channels: When does the feedback cost outweigh the spectral efficiency gain?" arXiv:1102.1552v1 [cs.IT].

[7] M. Mecking, "Resource allocation for fading multiple-access channels with partial channel state information," in *IEEE International Conference on Communications*, New York, NY, April 2002.

[8] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, vol. 46, no. 3, pp. 933–946, May 2000.