Analytic Solutions for Optimal Training on Fading Channels

Adam Panagos (Student) and Kurt Kosbar (Advisor)
Department of Electrical and Computer Engineering
University of Missouri - Rolla
Rolla, MO 65401
panagos@umr.edu, kosbar@umr.edu

ABSTRACT

Wireless communication systems may use training signals for the receiver to learn the fading coefficients of the channel. Obtaining channel state information (CSI) at the receiver is often times necessary for the receiver to correctly detect and demodulate transmitted symbols. The type of training signal, the length of time to spend training, and the frequency of training are all important parameters in these types of systems. In this work, we derive an analytic expression for calculating the optimal training parameters for continuously fading channels. We also provide simulation results showing why this training scheme is considered optimal.

KEY WORDS

Wireless Multiple-input multiple-output (MIMO), Training, Channel Capacity.

INTRODUCTION

Fundamental questions about training on multiple-input multiple-output channels were addressed by Hochwald *et al* in [1]. In this work, they examined training on a *block-fading channel*, i.e. a channel that is constant for a discrete time T, and changes to a new independent channel realization every T. They derived results that specified the optimal training signals to use, power allocation strategies, length of time to train, and the optimal training period.

This work was extended to a continuously fading channel model by Potter *et al* [2]. Using the concept of effective signal-to-noise ratio (ESNR), a lower bound on the channel capacity was derived, and the training parameters to maximize this lower bound were determined via Monte Carlo simulation. The simulation results were used to develop general guidelines for selecting optimal training parameters.

A proposition developed in [2], and re-stated in the following section, gives rules for how the optimal training parameters of the channel should change as the channel or system parameters change. While useful, this proposition was difficult to implement because it requires the original optimal parameters to be known. Optimal parameters for many different combinations of system parameters were tabulated, but we still desire a general expression for calculating values of the optimal training parameters directly. This work accomplishes this task.

The next section reviews previous results for optimal training on fading channels. Bounds and an approximation for the ergodic capacity of MIMO channels are presented next. These results are then used to reformulate the capacity lower bound from [2] to yield an analytic expression for optimal training parameters. Numerical results for an un-coded spatial multiplexing MIMO system or provided next. These results demonstrate why the training parameters calculated here are considered optimal.

PREVIOUS WORK

The parameters of the continuously varying MIMO channel investigated here are average signal-to-noise ratio, denoted ρ, channel fading rate, denoted f, number of transmit and receive antennas, denoted N_t and N_r respectively, and channel fading type, denoted β. Pure Rayleigh fading channels have $\beta = 0$, while a pure line-of-sight (LOS) Rician channel has $\beta = 1$. A mixture of the channel types can be obtained by selecting $0 < \beta < 1$.

The lower bound on channel capacity for the training scheme proposed in [2] is

$$C_{\text{lower}}(T, T_\tau) \triangleq \frac{1}{T} \sum_{t=1}^{T-T_\tau} \mathrm{E}_{\mathbf{H}} \left[\log_2 \det \left(\mathbf{I}_{N_r} + \frac{\rho_{\text{eff}}(t)}{N_t} \mathbf{H}\mathbf{H}^\dagger \right) \right], \tag{1}$$

where T is the training period, T_τ is the training length, $\rho_{\text{eff}}(t)$ is the time-varying ESNR, and $\mathbf{H}$ is the $N_t \times N_r$ channel matrix. More details on the channel, system parameters, and capacity lower bound can be found in [2].

The optimal training parameters, T^{opt} and T_τ^{opt}, are defined as the values of T and T_τ that maximize the channel capacity lower bound, i.e.

$$(T^{\text{opt}}, T_\tau^{\text{opt}}) = \arg \max_{T, T_\tau} C_{\text{lower}}(T, T_\tau). \tag{2}$$

The expectation operation ($\mathrm{E}[\cdot]$) of (1) is with respect to the random fading channel matrix $\mathbf{H}$. To evaluate (2), numerous Monte Carlo simulations were performed.

From these simulation results, guidelines for selecting optimal training parameters were summarized in the following proposition [2]:

Proposition 1 *Given a continuously varying MIMO channel defined by parameters* (β,ρ,f,N) *and corresponding optimal training parameters* $(T^{\text{opt}}, T_\tau^{\text{opt}})$*, optimal training parameters* $(\hat{T}^{\text{opt}},\hat{T}_\tau^{\text{opt}})$ *for the MIMO channel defined by parameters* $(\hat{\beta},\hat{\rho},\hat{f},\hat{N})$ *can be found by*

- *Decreasing (increasing)* T^{opt} *and* T_τ^{opt} *if* $\hat{\rho} - \rho > 0$ $(\hat{\rho} - \rho < 0)$
- *Decreasing (increasing)* T^{opt} *and* T_τ^{opt} *if* $\hat{f} - f > 0$ $(\hat{f} - f < 0)$
- *Decreasing (increasing)* T^{opt} *and* T_τ^{opt} *if* $\hat{N} - N < 0$ $(\hat{N} - N > 0)$
- *Decreasing (increasing)* T^{opt} *and* T_τ^{opt} *if* $\hat{\beta} - \beta < 0$ $(\hat{\beta} - \beta > 0)$

As mentioned in the previous section, while Proposition 1 provides general guidelines for how to change the training parameters if the optimal parameters are known, it does not allow the optimal training parameters to be calculated directly.

MIMO CHANNEL CAPACITY

We wish to directly calculate the optimal training parameters instead of determining them via lengthy Monte Carlo simulations. The key to accomplishing this is to eliminate the expectation operation of (1). For the following denote $\alpha = \frac{K}{K+1}$ and $\beta = \frac{1}{K+1}$ as parameters related to the LOS strength of the fading channel (often called the Rician K-factor). Also, while simulation results of [2] only considered systems with the same number of transmit and receive antennas, the following results allow N_t transmit and N_r receive antennas. We also let $n = \min\{N_t, N_r\}$ and $m = \max\{N_t, N_r\}$. These results include Rayleigh fading as a special case by letting $K = 0$ $(\alpha = 0, \beta = 1)$. For proofs of these results refer to [3].

Theorem 2 (Upper Bound) *For a rank-1 Rician MIMO channel without transmit or receive correlation,* N_t *transmit antennas,* N_r *receive antennas, and average signal-to-noise ratio of* ρ*, we have an ergodic capacity upper bound of*

$$\begin{aligned} C &\leq \log_2(1 + \rho\tilde{c}_1 + \rho\tilde{c}_2) + (n-1)\log_2(1 + \rho\tilde{c}_1) \\ &\triangleq C_{\text{upper}}(N_t, N_r, \rho, \alpha) \end{aligned} \tag{3}$$

where $\tilde{c}_1 = \beta$, $\tilde{c}_2 = n\alpha$ for $N_r \leq N_t$ and $\tilde{c}_1 = \frac{\beta m}{n}$, $\tilde{c}_2 = m\alpha$ for $N_r > N_t$.

Theorem 3 (Lower Bound) *For a rank-1 Rician MIMO channel without transmit or receive correlation,* N_t *transmit antennas,* N_r *receive antennas, and average signal-to-noise ratio of* ρ*, we have an ergodic capacity lower bound of*

$$C \geq n\log_2\left(1 + \frac{\rho\beta}{n}\right) \triangleq C_{\text{lower}}(N_t, N_r, \rho, \alpha). \tag{4}$$

Proposition 4 (Rician Channel Approximation) *The capacity of a rank-1 Rician MIMO channel without correlation is well approximated by*

$$C \approx k_1 C_{\text{upper}}(N_t, N_r, \rho, \alpha) + k_2 C_{\text{lower}}(N_t, N_r, \rho, \alpha) \triangleq C_{\text{approx}}(N_t, N_r, \rho, \alpha) \tag{5}$$

over the SNR range 0 to 20 dB for a set of constants $k_1, k_2 \in \mathbb{R}$.

The constants k_1 and k_2 required by Proposition 4 have been tabulated in the appendix for Raleigh Fading channels (i.e. $K = 0$). The error between the approximation and true channel capacity has also been tabulated. A comparison of the approximation and true channel capacity can be seen in Figures 1 and 2. These figures and tables show that the capacity approximation of Proposition 4 is quite good.

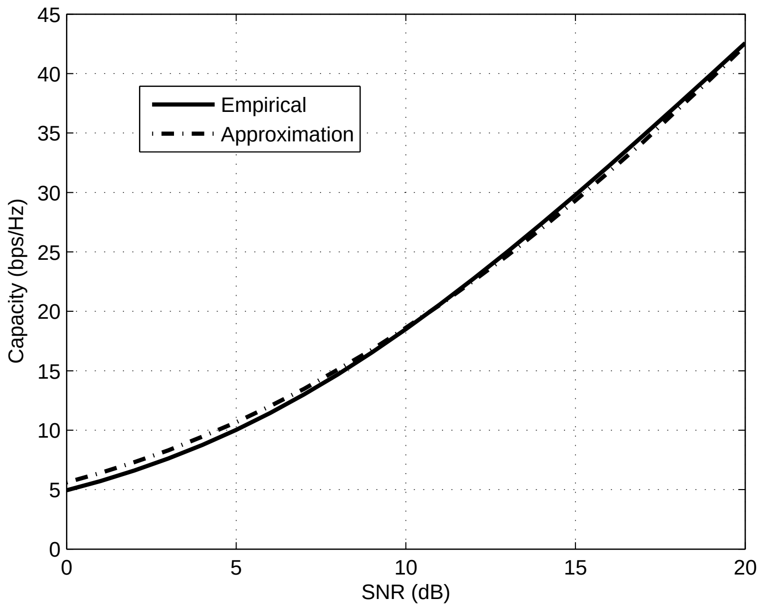

Figure 1: Capacity and approximation for $M = N = 10$ antennas with $\beta = 0.25$. Average mean absolute error is 3.75 percent. Even for this "worst-case" scenario, the approximation is quite good.

For a general MIMO system with N_t transmit antennas, N_r receive antennas, average SNR ρ, and random channel matrix $\mathbf{H}$, the ergodic channel capacity is the well-known equation [4, 5, 6]

$$\mathrm{E}_{\mathbf{H}} \left[\log_2 \det \left(\mathbf{I}_{N_r} + \frac{\rho}{N_t} \mathbf{H}\mathbf{H}^\dagger \right) \right]. \tag{6}$$

Combining this with Proposition 4, the lower bound of (1) can be re-written as

$$\widehat{C}_{\text{lower}}(T, T_\tau) \triangleq \frac{1}{T} \sum_{t=1}^{T-T_\tau} C_{\text{approx}}(N_t, N_r, \rho_{\text{eff}}(t), \alpha). \tag{7}$$

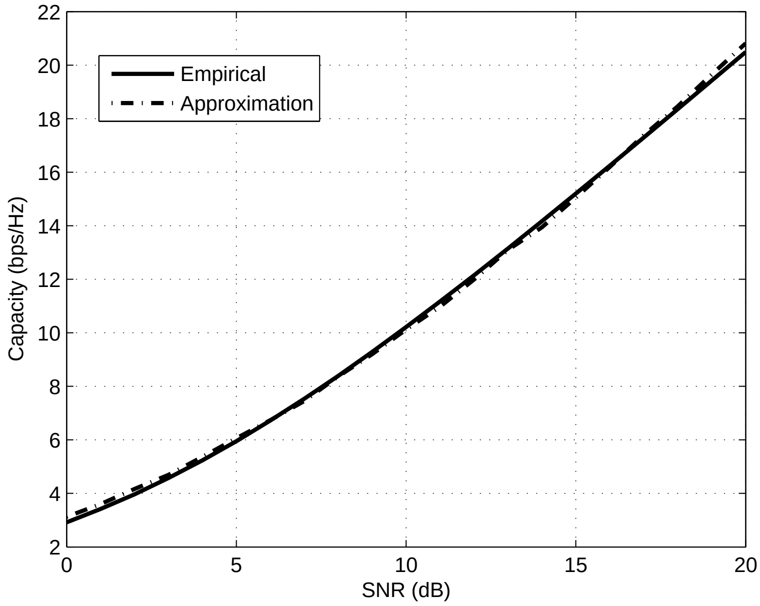

Figure 2: Capacity and approximation for $M = N = 4$ antennas with $\beta = 0.50$. This is a more "typical" scenario with average mean absolute error of 1.79 percent.

The optimal training parameters can now be calculated by evaluating

$$(T^{\text{opt}}, T_\tau^{\text{opt}}) = \arg \max_{T, T_\tau} \widehat{C}_{\text{lower}}(T, T_\tau) \tag{8}$$

directly, without any need for lengthy simulations. In Table 4 of the appendix we have tabulated the optimal training parameters calculated analytically via (8) and the optimal training parameters as determined vie Monte Carlo simulation and (2) for a Rayleigh Fading environment with $N_t = N_r = 2$. Note the similarity in results as desired.

NUMERICAL EXAMPLES

In the previous section we derived an analytic expression for calculating optimal training parameters for fading channels. The results comparison of Table 4 confirm the validity of this approach. In this section we investigate why these parameters are considered optimal. The next sub-section describes the spatial multiplexing system that was simulated, followed by performance results that demonstrate why these are denoted *optimal* training parameters.

- Spatial Multiplexing System

We simulated an un-coded spatial multiplexing system. A block diagram of the system for the case of $N_t = N_r = 2$ can be seen in Figure 3.

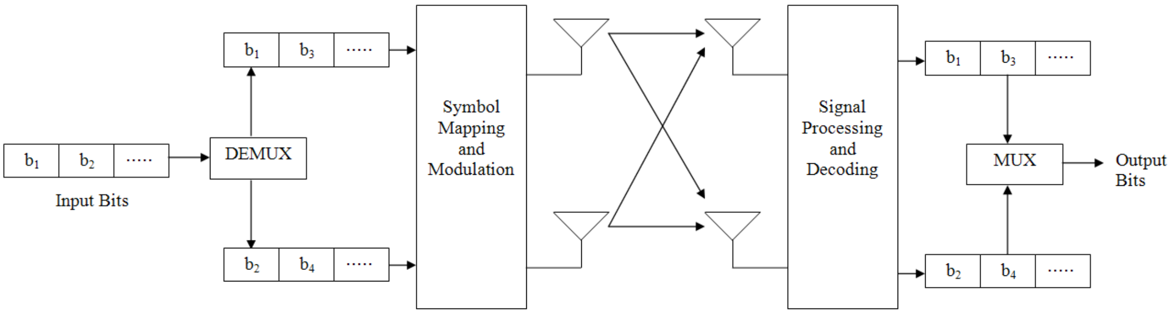

Figure 3: Spatial Multiplexing System Block Diagram for 2×2 System.

A bitstream is multiplexed to each antenna, where bit-to-symbol mapping is performed. A QAM constellation was used for each antenna. Jakes fading channel model was used to construct the time-varying channel matrix $\mathbf{H}(t)$ [7, 8]. At time instance t, the data received at antenna k is

$$y_{t,k} = \sqrt{\frac{\rho}{N_t}} \sum_{r=1}^{N_t} h_{r,k,t} s_{t,r} + v_{t,k}, \tag{9}$$

where $h_{r,k,t}$ are entries of the $N_t \times N_r$ channel matrix $\mathbf{H}(t)$ at time t, $s_{t,r}$ are entries of the $T \times N_t$ transmitted signal matrix $\mathbf{S}$, and the independent identically distributed complex Gaussian noise has variance of $1/2$ per dimension (i.e. $v_{t,n} \sim \mathcal{CN}(0, 1)$. We assume that $\mathbf{H}$ and $\mathbf{S}$ are normalized such that $E\left[||\mathbf{H}||^2\right] = N_t N_r$ and $E\left[||\mathbf{S}||^2\right] = T N_t$, so that ρ represents the average SNR at the receive antennas.

In our communication system, each block of T symbol times is divided into two distinct transmission phases: a training phase that lasts for T_τ symbol times, and a data transmission phase that lasts for $T - T_\tau$ symbol times. Thus, the transmitted and received signals can be partitioned as

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{\text{train}} \\ \mathbf{S}_{\text{data}} \end{pmatrix} \text{ and } \mathbf{Y} = \begin{pmatrix} \mathbf{Y}_{\text{train}} \\ \mathbf{Y}_{\text{data}} \end{pmatrix}, \tag{10}$$

where $\mathbf{S}_{\text{train}}$ is $T_\tau \times N_t$, $\mathbf{Y}_{\text{train}}$ is $T_\tau \times N_r$, $\mathbf{S}_{\text{data}}$ is $T - T_\tau \times N_t$, and $\mathbf{Y}_{\text{data}}$ is $T - T_\tau \times N_r$.

For simplicity at the receiver, we ignore the time-varying nature of the channel and assume the channel is constant over T symbol times. During channel training, signals known to both the transmitter and receiver are transmitted for T_τ symbol times. The receiver can then calculate the maximum likelihood (ML) estimate $\hat{\mathbf{H}}$ of the channel as [7]

$$\hat{\mathbf{H}} = \sqrt{\frac{N_t}{\rho}} (\mathbf{S}_{\text{train}}^\dagger \mathbf{S}_{\text{train}}) \mathbf{S}_{\text{train}}^\dagger \mathbf{Y}_{\text{train}}. \tag{11}$$

Our training signals are constructed as

$$\mathbf{S}_{\text{train}} = \sqrt{T_\tau} \cdot \text{orth}(\mathbf{A}), \tag{12}$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_{N_t} \\ \mathbf{1}_{(T_\tau - N_t, N_t)} \end{pmatrix}, \tag{13}$$

$\text{orth}(\cdot)$ is an orthonormal basis for the range of $\mathbf{A}$, and $\mathbf{1}_{M,N}$ is an $M \times N$ matrix of all ones. This construction ensures that

$$\mathbf{S}_{\text{train}}^\dagger \mathbf{S}_{\text{train}} = T_\tau \cdot \mathbf{I}_{N_t}, \tag{14}$$

which is shown to be optimal for the block-fading channel in [1].

For demodulation at the receiver, we ignore the time varying nature of the channel and assume the ML estimate of the channel is correct. Thus, during the data transmission phase, the received signal is approximately

$$\mathbf{Y}_{\text{data}} \approx \sqrt{\frac{\rho}{N_t}} \mathbf{S}_{\text{data}} \hat{\mathbf{H}} + \mathbf{V}_{\text{data}}. \tag{15}$$

We decode the transmitted symbols as

$$\mathbf{S}_{\text{data}} = q\left(\sqrt{\frac{N_t}{\rho}} \mathbf{Y}_{\text{data}} \hat{\mathbf{H}}^{-1}\right), \tag{16}$$

where $q(\cdot)$ is a slicing function that rounds each element of $\mathbf{S}_{\text{data}}$ to the nearest symbol of the underlying QAM constellation. This demodulation rule is clearly sub-optimal, but our goal here is to show how the training parameters *affect* system performance, not how to optimally demodulate signals.

- Results

Figure 4 shows the symbol-error rate (SER) performance of the spatial multiplexing system for parameters $\rho = 10$ dB, $N_t = N_r = 3$, $f = 0.004$, $\beta = 1$, and different combinations of training parameters T and T_τ. The rate is calculated as $R = \frac{T - T_\tau}{T}$ for each data point. The filled "•" is the performance achieved by the optimal training parameters of $(T^{\text{opt}}, T_\tau^{\text{opt}}) = (23, 4)$ calculated from (8). Other training parameters were simulated and are plotted as "x". We see that an increase in rate from the optimal operating point leads to significant increases in SER, while only a marginal decrease in SER can be achieved for a significant decrease in rate. The solid line separating the obtainable and unobtainable regions was obtained from the convex hull of the data points. The convex hull of a set of points is the minimal convex set containing the data points.

Similar results are shown in Figure 5. Results shown there used $N_t = N_r = 3$, $f = 0.004$, $\beta = 0.50$, and various ρ. Only the convex hull boundary and optimal solutions (represented by a bullet, "•") are shown. As before, the optimal training solution is at a location that jointly maximizes the system rate and minimizes the SER. It is for this reason these training parameter values are denoted optimal. Results for other system parameters not included here are similar.

Figure 6 shows the error distribution of the spatial multiplexing system and emphasizes the sub-optimal detection technique used here. The parameters used for this plot were $T = 30, T_\tau = 4, M = 3, N = 3$, and $\rho = 10$ dB. The channel parameter f was varied from $0.002, 0.004, 0.008$. For these training parameters, there are $T - T_\tau = 26$ data symbols per block. Thus, errors can happen in one of 26 data slots. Due to the continuous variation of the channel, the static estimate of the channel used for demodulation becomes worse as times evolves. Thus, the probability of making an error increases with time. Also, as expected, the probability of a symbol error increases as the channel time variation increases.

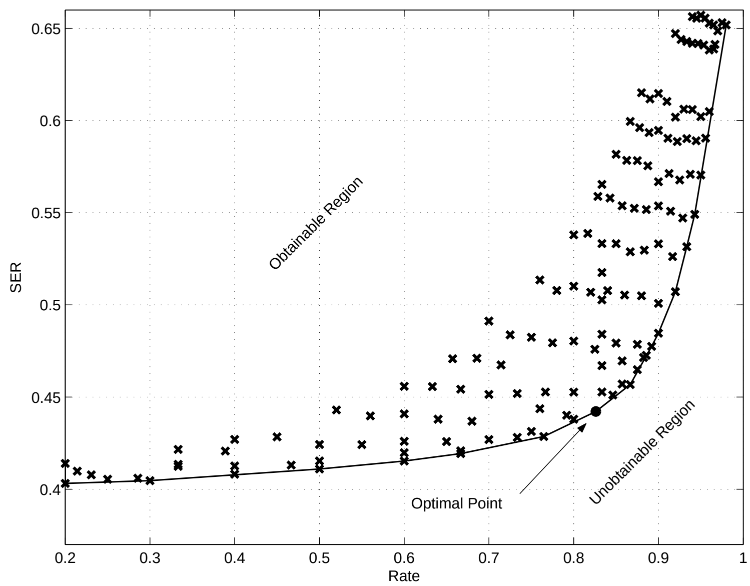

Figure 4: Symbol error rate (SER) performance for different training parameters.

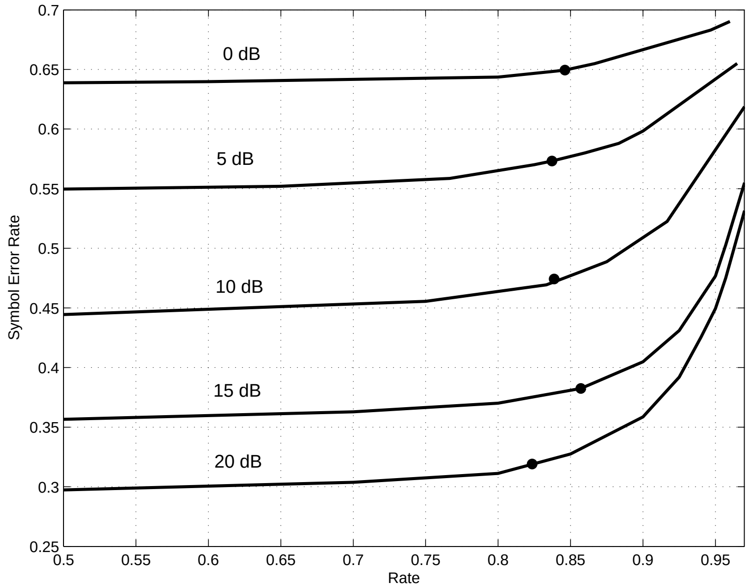

Figure 5: Convex hull of rate and SER for $\beta = 0.50$, $M = N = 3$, $f = 0.004$, and various SNR.

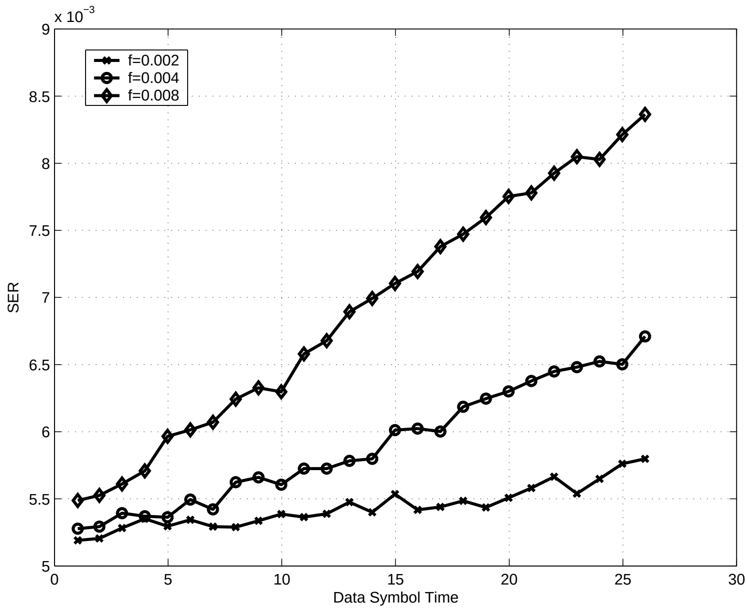

Figure 6: Error distribution in time.

CONCLUSIONS

Previous work on training for continuously varying fading channels used extensive simulations to developed rough guidelines for selecting optimal training parameters. In this work, we provided an analytic expression for optimal training parameters that can be calculated directly and easily. We have provided simulation results for a simple spatial multiplexing system that shows why these results are considered optimal. The optimal training parameters jointly minimize the system SER while maximizing system throughput.

APPENDIX

Table 1: Capacity Approximation Parameter k_1 for $K = 0, (\beta = 1)$

$N_t \setminus N_r$	1	2	3	4	5	6	7	8	9	10
1	0.30	0.82	0.84	0.93	0.94	0.94	0.96	0.97	0.92	0.92
2	0.72	0.69	0.81	0.83	0.86	0.89	0.91	0.93	0.95	0.95
3	0.59	0.81	0.80	0.78	0.83	0.86	0.88	0.90	0.91	0.92
4	0.28	0.78	0.79	0.78	0.81	0.82	0.84	0.87	0.88	0.89
5	0.97	0.86	0.82	0.82	0.78	0.79	0.82	0.83	0.85	0.87
6	0.29	0.83	0.85	0.82	0.81	0.79	0.79	0.82	0.83	0.85
7	0.27	0.90	0.87	0.84	0.81	0.80	0.78	0.79	0.81	0.82
8	0.84	0.89	0.89	0.86	0.83	0.81	0.80	0.78	0.79	0.80
9	0.08	0.91	0.90	0.87	0.85	0.83	0.81	0.79	0.79	0.79
10	0.41	0.93	0.90	0.89	0.86	0.84	0.83	0.81	0.79	0.78

Table 2: Capacity Approximation Parameter k_2 for $K = 0, (\beta = 1)$

$N_t \setminus N_r$	1	2	3	4	5	6	7	8	9	10
1	0.55	0.14	0.16	0.06	0.06	0.07	0.04	0.03	0.02	0.02
2	0.21	0.16	0.12	0.15	0.14	0.11	0.10	0.07	0.05	0.05
3	0.36	0.09	0.01	0.15	0.13	0.12	0.11	0.10	0.09	0.09
4	0.68	0.18	0.12	0.04	0.09	0.13	0.14	0.11	0.12	0.12
5	0	0.10	0.12	0.06	0.05	0.12	0.12	0.15	0.14	0.12
6	0.68	0.15	0.10	0.11	0.07	0.03	0.12	0.11	0.13	0.13
7	0.71	0.07	0.09	0.11	0.12	0.09	0.05	0.11	0.13	0.15
8	0.14	0.09	0.08	0.10	0.12	0.11	0.08	0.06	0.10	0.14
9	0.90	0.07	0.07	0.10	0.11	0.11	0.10	0.10	0.03	0.10
10	0.57	0.05	0.08	0.08	0.11	0.12	0.10	0.10	0.10	0.05

Table 3: Capacity Approximation Percent Error for $K = 0, (\beta = 1)$

$N_t \setminus N_r$	1	2	3	4	5	6	7	8	9	10
1	3.01	1.14	0.72	0.46	0.52	0.46	0.19	0.33	0.33	0.20
2	1.60	1.86	0.80	0.68	0.36	0.42	0.30	0.28	0.34	0.25
3	1.21	1.47	1.73	1.00	0.67	0.38	0.33	0.38	0.36	0.31
4	0.87	0.77	1.47	1.59	0.97	0.67	0.42	0.48	0.28	0.20
5	0.83	0.92	0.94	1.24	1.74	1.20	0.74	0.51	0.45	0.31
6	0.81	0.59	0.78	0.80	1.26	1.51	1.00	0.85	0.57	0.42
7	0.64	0.44	0.67	0.73	1.03	1.20	1.40	1.27	0.90	0.65
8	0.79	0.44	0.60	0.73	0.82	1.21	1.28	1.47	1.23	0.90
9	0.61	0.44	0.45	0.69	0.79	0.98	1.10	1.31	1.52	1.22
10	0.66	0.27	0.52	0.42	0.69	0.84	0.91	1.15	1.33	1.42

Table 4: Rayleigh Fading Optimal Training Parameters ($N_t = N_r = 2$)

SNR dB	f	T^{sim}	T_τ^{sim}	T^{analytic}	$T_\tau^{\mathrm{analytic}}$
0	0.002	59	8	56	8
5	0.002	48	6	45	6
10	0.002	37	4	33	4
15	0.002	24	3	25	3
20	0.002	18	2	17	2
0	0.004	34	6	34	6
5	0.004	26	4	28	5
10	0.004	20	3	20	3
15	0.004	14	2	14	2
20	0.004	9	1	12	2
0	0.008	19	4	19	4
5	0.008	14	3	16	3
10	0.008	12	2	12	2
15	0.008	11	2	10	2
20	0.008	6	1	9	2

References

[1] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links," *IEEE Trans. Inform. Theory*, vol. 49, pp. 951–963, Apr. 2003.

[2] C. Potter, A. Panagos, K. Kosbar, and W. Weeks, "Optimal training parameters for continuously varying mimo channels," *Proc. International Telemetry Conference (ITC'05)*, Oct. 2005.

[3] A. Panagos and K. Kosbar, "On mimo rician channel capacity," *Submitted to Wireless Personal Communications*, Dec. 2005.

[4] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[5] E. Telatar, "Capacity of multi-antenna gaussian channels," *AT&T Bell Laboratories Internal Tech. Memo.*, June 1995.

[6] D. Gesbert *et al.*, "From theory to practice: An overview of mimo space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[7] C. Peel and A. Swindlehurst, "Effective snr for space-time modulation over a time-varying rician channel," *submitted to IEEE Transaction on Wireless Communicatinos*, vol. 52, no. 1, pp. 17–23, Jan. 2004.

[8] W. Jakes, *Microwave Mobile Communications*. Wiley-IEEE, 1974.