OPTIMAL TRAINING PARAMETERS FOR CONTINUOUSLY VARYING MIMO CHANNELS

Christopher G. Potter, Adam G. Panagos, Kurt Kosbar, and William Weeks

Telemetry Learning Center
Department of Electrical and Computer Engineering
University of Missouri – Rolla
Rolla, MO 65409-0040

ABSTRACT

To correctly demodulate a signal sent through a multiple-input multiple-output (MIMO) channel, a receiver may use training to learn the channel parameters. The choice of training parameters can significantly impact system performance. Training too often yields low throughput while training infrequently produces poor channel estimates and increased transmission errors. Previous work on optimal training parameters has focused on the block fading Rayleigh model. This work examines a more general case; finding the training parameters that maximize throughput for a continuously varying channel. Training parameters that maximize a lower bound on channel capacity are determined via simulation, and general guidelines are presented for selecting optimal training parameters.

KEYWORDS

MIMO Channels, Optimal Training, Channel Capacity, Rayleigh Channel, Rician Channel.

INTRODUCTION

Theoretical results show that under certain conditions, use of multiple transmit and receive antennas in wireless telemetry applications can significantly increase spectral efficiency [1-3]. The promise of increased spectral efficiency is made under the assumption that the receiver knows the channel conditions perfectly. Due to noise and other undesirable real-world conditions, perfect knowledge of the channel is never realized in practice. One possible solution for obtaining channel state information (CSI) is to transmit a known signal (i.e. a training sequence) from the transmitter and perform estimation at the receiver.

The time spent transmitting the training sequence cannot be used to transmit data, thus lowering the effective throughput. On the other hand, too little time spent training results in poor channel estimates and increased transmission errors. Thus the frequency and amount of time spent training has a significant impact on system performance. In this work we present guidelines for training on a continuously varying MIMO channel.

The next section describes previous work for optimal training on block fading MIMO channels. These results are extended to continuously fading channels in the following section. A summary of the simulations performed to determine optimal training parameters for different continuously varying channels is presented, along with guidelines for selecting the optimum training period and training length. The work completed here and ideas currently being explored are summarized in the conclusion.

TRAINING ON THE BLOCK FADING MIMO CHANNEL

Multiple-input multiple-output communication systems have higher spectral efficiency than comparable single-input single-output communication systems. Since much of the literature regarding MIMO communication systems assumes the receiver knows the channel perfectly, it is essential that the receiver be able to estimate the channel matrix elements accurately. Without an accurate estimate of the channel, the receiver will not be able to demodulate data correctly.

Two pilot training systems [4,5] have recently been proposed that allow estimation of the channel matrix elements to be performed at the receiver. These pilot training systems assume a block-fading environment which means that the channel is constant for a discrete time interval T, after which it changes to a new and independent channel realization. These pilot schemes dedicate a fraction of the first T time instants to sending known data symbols. Transmission of the training sequence allows the receiver to form an estimate of the channel matrix elements and demodulate the subsequent data correctly. These works also develop theorems that specify the type of training signals to use for MIMO channel training, the optimal fraction of total transmit power to assign to the training period, and the fraction of total block length time T to spend training. Before addressing our new results for the continuously fading channel, we review the relevant results from these works.

The block fading MIMO channel model is described by [6]

$$X = \sqrt{\frac{\rho}{M}} SH + V \qquad (1)$$

where X is the received $T \times N$ matrix, ρ is the average SNR at each receiver antenna, M is the number of transmit antennas, N is the number of receive antennas, S is the $T \times M$ transmitted signal whose entries have unity average mean square value. The entries of the $M \times N$ channel matrix H are independent identically distributed circularly symmetric complex Gaussian random variables with variance of $1/2$ per dimension. The $T \times N$ noise matrix V consists of i.i.d. complex Gaussian random variables.

Consider the following training scheme for the MIMO block fading channel. During each T symbol block of time, the first T_τ symbols times are spent transmitting a known training sequence. When the power used during the training phase and data transmission phase are equal, and unitary training signals are used for training, a lower bound on the channel capacity is [4 equation (40)]:

$$C(T_\tau) \geq \frac{T - T_\tau}{T} \mathrm{E}\left[\log\det\left(I_M + \frac{\rho_{eff}}{M} HH^\dagger \right) \right] \tag{2}$$

where

$$\rho_{eff} = \frac{\rho^2 T_\tau / M}{1 + (1 + T_\tau / M)\rho} \tag{3}$$

Note that for finite T_τ we have $\rho_{eff} < \rho$. The effective SNR takes into account the error made estimating the channel due to using a finite amount of training data. The error between the channel estimate and the true channel can be treated as additional additive noise, which in the worst case is also complex additive Gaussian noise. Thus, the authors of [4] establish the lower bound of Equation (2). Note that we have

$$\lim_{T_\tau \to \infty} \rho_{eff} = \lim_{T_\tau \to \infty} \frac{\rho^2 T_\tau / M}{1 + (1 + T_\tau / M)\rho} = \frac{\rho^2 / M}{\rho / M} = \rho \tag{4}$$

As desired, the error between the channel estimate and the true channel can be made as small as necessary by training for a sufficiently long period of time, as long as the block-length of the channel T is adequately long. This results in an effective SNR which is equal to the SNR due to just additive noise, i.e. $\rho_{eff} = \rho$.

Note the scale factor of $\frac{T - T_\tau}{T}$ in Equation (2). This scale factor accounts for the portion of time spent sending the training sequence that does not result in any data being sent. This scale factor suggests that choosing T_τ small will maximize the lower bound of Equation (2). However, choosing a small T_τ decreases the effective SNR and hence decreases the $\det(\cdot)$ term of the expression. Thus, choosing the optimal value of T_τ is not trivial. For certain cases, closed-form results do exist. For instance, at low SNR the optimal choice is $T_\tau = T/2$ and at high SNR the optimal choice is $T_\tau = M$. For an intuitive explanation of these results in the low and high SNR regimes see [4] Section III-D. For other scenarios, the optimal choice for T_τ can be found via simulation.

TRAINING ON THE CONTINOUSLY FADING MIMO CHANNEL

We propose using the training scheme for block-fading channel on the continuously fading channel. That is, transmission time will be divided into blocks of T symbol times, the first

T_τ symbol times will be used to send a training sequence to form an estimate of the channel at the receiver, and the remaining $T-T_\tau$ symbol times will be used to transmit data. While this method is attractive due to its simplicity, the continuous nature of the channel causes static estimates formed during the training phase to become worse as time evolves. However, using the concept of effective SNR, it is still possible to consider the error between the estimate of the channel and the actual channel as additive noise. The fundamental difference is instead of a constant estimation error over T symbol periods as in the block-fading channel scenario, the continuously varying channel estimation error will increase with time, causing the effective SNR to decrease as a function of time.

The continuously fading channel also presents the new challenge of choosing the optimal training period T. The block-fading channel provided a logical choice of using a training period equal to the block-fading length. There is no obvious choice for T on the time-varying channel. Thus, both the optimum training time, i.e. T_τ, and the optimal training period T must now be determined.

Effective SNR for time-varying MIMO channels has been investigated previously [6]. The channel model used encompasses both diffuse (i.e. Rayleigh) and specular (i.e. Rician) channels and is defined by

$$H(t)=\sqrt{1-\beta}H^S+\sqrt{\beta}H^D(t) \tag{5}$$

The parameter $0\le\beta\le1$ controls what type of a channel is being used. Setting $\beta=0$ results in a pure Rician channel while $\beta=1$ gives a pure Rayleigh channel. The specular portion of the channel is rank-1, assumed known at the receiver, and is not time varying. The diffuse portion of the channel is time varying and thus indicated with the (t) notation. A first order auto-regressive model is used to describe the variation of the diffuse channel with symbol time

$$H^D(t+r)=\sqrt{\alpha(t)}H^D(r)+\sqrt{1-\alpha(t)}W(r+t) \tag{6}$$

where $H^D(r)$ is the estimate of the diffuse portion of the channel at reference time r and $H^D(t+r)$ is the value of the diffuse channel component t symbol times later. The entries of the complex additive Gaussian noise matrix $W(r+t)$ are independent and are also independent in sample time. The function $\alpha(t)$ determines the time-varying nature of the channel. The choice made here is based on Jakes model [7]

$$\alpha(t)=\left(J_0(2\pi tf)\right)^2 \tag{7}$$

The channel variation parameter f determines how quickly the channel changes with time and is defined as $f=f_dT_s$ where f_d is the maximum Doppler frequency and T_s is the sampling period. As with the block-fading channel model the error between the channel estimate and true channel value is treated as additional noise. This allows us to treat the channel as constant but the channel SNR as a function of time. With this channel model, the time-varying effective SNR is given as [6]

$$\rho_{eff}(t) = \frac{\rho(1-\beta) + \rho\beta\alpha(t)}{1 + \alpha(t)\frac{M}{T_\tau} + (1-\alpha(t))\rho\beta} \tag{8}$$

Since channel capacity is a function of SNR, and the effective SNR is now a decreasing function of time, channel capacity is also a decreasing function of time. For the continuously varying channel, the lower bound on capacity that we wish to maximize is thus

$$C(T, T_\tau) \geq \frac{1}{T}\sum_{t=1}^{T-T_\tau} \mathrm{E}\left[\log\det\left(I_M + \frac{\rho_{eff}(t)}{M} HH^\dagger\right)\right] \tag{9}$$

As previously mentioned, this lower bound now must be maximized over two variables, T and T_τ. Note the channel and system parameters that effect this expression. The channel matrix H is an $M \times N$ matrix, $\rho_{eff}(t)$ depends on the channel type β, the rate of change of the channel f, the channel SNR ρ, the amount of time spent training T_τ, and the sum itself is a function of T and T_τ. The optimal values of T and T_τ that maximize this lower bound and how these values are effected by the various parameters have been determined via simulation. The results of the simulations and general guidelines for choosing optimal training parameters are presented in the following section.

OPTIMAL TRAINING PARAMETERS AND GUIDELINES

Simulations were performed to determine the values of T and T_τ that maximize the lower bound of Equation (9). The values that maximize the lower bound are called the optimal training parameters. These optimal parameters were found for different channel types ($\beta = 0.25, 0.50, 0.75,$ and 1), signal to noise ratios ($\rho = 0, 5, 10, 15,$ and 20 dB), different numbers of transmit and receive antennas ($M, N = 2, 3, 4$), and different channel variations ($f = 0.002, 0.004, 0.008$).

A plot of the maximum lower bound as a function of channel variation has been plotted in Figure 1. The curves corresponding to $f = 0.002, 0.004,$ and 0.008 were generated by finding the optimal training parameters to maximize Equation (9). The training period that maximized this expression was then used in Equation (2) to see what the corresponding upper bound would be for the corresponding block-fading channel. As expected, an increase in channel variation leads to a decrease in the lower bound. This figure also shows the lower bound of Equation (9) derived in this work agrees with the previously derived block fading lower bound of Equation (2).

A plot of the maximum lower bound as a function of channel variation has been plotted in Figure 1. The curves corresponding to $f = 0.002, 0.004,$ and 0.008 were generated by finding the optimal training parameters to maximize Equation (9). The training period that maximized this expression was then used in Equation (2) to see what the upper bound would be for the corresponding block-fading channel. As expected, in increase in channel variation leads to a decrease in the lower bound. This figure also shows the lower bound of Equation (9) derived in this work agrees with the previously derived block fading lower bound of Equation (2).

A plot of the optimal block length versus effective SNR with curves corresponding to values of f = .002, .004, and .008 can be seen in Figure 2. To gain insight into this result, recall that the correlation of the channel coefficients has been modeled by Equation (7). This equation has been plotted for the same values of f that were chosen for the simulation and can be seen in Figure 3. Clearly, as f increases, the autocorrelation of the diffuse component of the channel decreases with respect to time. This effect is undesirable since it will decrease the amount of time we can rely on our channel estimate, which is known as the *coherence interval.* A smaller coherence interval should result in a decrease in the optimal block length which is apparent in Figure 2.

For Figure 4, optimal training length T_τ versus SNR was plotted with the parameters M, N, and f being varied. A couple of observations can be made. The first is that as the number of antennas increase, the receiver must spend more time training because there are more channel coefficients to estimate. This setback is amplified at low SNR because the noise has a more significant effect on corrupting the channel coefficients. This reasoning can be justified mathematically by observing (8), and (9). The second term in the denominator of (8) is proportional to M, thus reducing its value. Also, the second term in the determinant of (9) is proportional to $\rho_{eff}(t)$ and inversely proportional to M, which results in a decrease of this second term, which lowers the value of (9).

The second observation is that as f decreases, T_τ increases. Since the diffuse component of the channel is more correlated in time when f is small (i.e. the channel is not varying much over time), it is better for the receiver to train longer to better approximate the channel coefficients. Through this same reasoning, if the diffuse channel is less correlated in time, it is more beneficial for the receiver to spend less time trying to estimate the channel coefficients and spend more time transmitting data. This result can be verified mathematically by observing both Figure 3 and Equation 8. Recall that as f decreases, $\alpha(t)$ does not decay as rapidly. Therefore to maximize $\rho_{eff}(t)$ which in turn maximizes the capacity, the focus is placed on the second term in the denominator of (8). Clearly, an increase in the training interval T_τ is beneficial to maximize $\rho_{eff}(t)$.

Figure 5 was constructed by observing SNR versus maximum data rate while varying M, N, and β. The results from this plot are intuitive. As β increases, the channel is dominated by a non-line-of-sight component. For this reason, we are not able to send as much throughput compared to low values of β. This observation can be verified by observing the following. It is obvious that

$$\lim_{\beta \to 0} \rho_{eff}(t) = \frac{\rho}{1 + \frac{\alpha(t) \cdot M}{T_\tau}} \quad (10)$$

$$\lim_{\beta \to 1} \rho_{eff}(t) = \frac{\rho}{\frac{1+\rho}{\alpha(t)} + \frac{M}{T_\tau} - \rho} \quad (11)$$

Since $0 \le \alpha(t) \le 1$ and for the most part $\frac{M}{T_\tau} \le 1$ is valid, the value of (10) will be approximately ρ. More importantly one should observe that a small value of $\alpha(t)$ does not have a negative effect on (10) which in turn means there is no negative effect on (9) and subsequently the maximum data rate. However, looking at (11) clearly indicates that if $\alpha(t) \to 0$, then $\frac{1+\rho}{\alpha(t)} \to \infty$ which implies $\lim_{\beta \to 1} \rho_{eff}(t) = 0$. This will have a negative effect on Equation 9 and subsequently the maximum data rate.

From these results we propose the following guidelines for choosing optimal training parameters.

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

1. Decreasing (increasing) T^{opt} and T_τ^{opt} if $\hat{\rho} - \rho > 0$ ($\hat{\rho} - \rho < 0$).
2. Decreasing (increasing) T^{opt} and T_τ^{opt} if $\hat{f} - f > 0$ ($\hat{f} - f < 0$).
3. Decreasing (increasing) T^{opt} and T_τ^{opt} if $\hat{N} - N < 0$ ($\hat{N} - N > 0$).
4. Decreasing (increasing) T^{opt} and T_τ^{opt} if $\hat{\beta} - \beta < 0$ ($\hat{\beta} - \beta > 0$).

If just one channel parameter is changed, Proposition 1 makes it simple to decide how to change the training parameters to maximize the lower bound on capacity. If more than one channel parameter is changed, it is possible that Proposition 1 suggests to both increase *and* decrease the training parameters. In this case, the figures in the appendix can be used to determine which channel parameter most significantly affects the training parameters. For example, at low SNR, changing from $\beta = 0.25$ to $\hat{\beta} = 1$ will have a much more significant impact on the optimal training parameters compared to changing from $N = 2$ to $\hat{N} = 3$. In this case, the change in antenna elements should be ignored, $\hat{\beta} - \beta > 0.75 > 0$, and the training parameters should be decreased.

CONCLUSION

Previous work on optimal training parameters for a block fading MIMO channel has been extended to a continuously varying MIMO channel. The channel model used encompasses both Rayleigh and Rician channel types. A specific training scheme for the continuously varying channel was proposed and using the concept of effective signal-to-noise ratio a lower bound on the channel capacity was derived. The training parameters that maximized the lower bound were determined via simulation for a variety of different channel parameters, and general guidelines for selecting the optimal training parameters were provided in Proposition 1. These guidelines suggest that both the optimal training period and training time increase as channel SNR decreases, increase as the number of transmit and receive antennas is increased, and increase as the rate of change of the channel decreases. The authors are currently pursuing analytic approximations for the channel capacity lower bound that will allow optimal training parameters to be calculated directly instead of determined via simulation.

REFERENCES

[1] Gesbert, D., Shafi, M., Da-shan, S., Smith, P.J., Naguib, A., "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems", IEEE Journal on Selected Areas in Communications, vol. 21, Apr., 2003, pp. 281-302.

[2] Foschini, G.J., Gans, M.J., "On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas", Wireless Personal Communications, no. 6, 2003, pp. 311-335.

[3] Telatar, I.E., "Capacity of Multiantenna Gaussian Channels", AT&T Bell Labs Internal Technical Memo, Jun., 1995.

[4] Hassibi, B. and Hochwald, B.M., "How Much Training is Needed in Multiple Antenna Wireless Links?", IEEE Transactions on Information Theory, vol. 49, Apr., 2003, pp. 951-963.

[5] Samardzija, D. and Mandayam, N., "Pilot-Assisted Estimation of MIMO Fading Channel Response and Achievable Data Rates", IEEE Transactions on Signal Processing, vol. 51, Nov., 2003, pp. 2882-2890.

[6] Peel, C.B. and Swindlehurst, A.L., "Effective SNR for Space-Time Modulation Over a Time-Varying Rician Channel", IEEE Transactions on Communications, vol. 52, Jan., 2004, pp. 17-23.

[7] Jakes, W.C., "Microwave Mobile Communications", 1974, Wiley.

APPENDIX

Figure 1 – Rayleigh Fading: Maximal Lower Bound vs. SNR as a function of 'f': M,N = 2

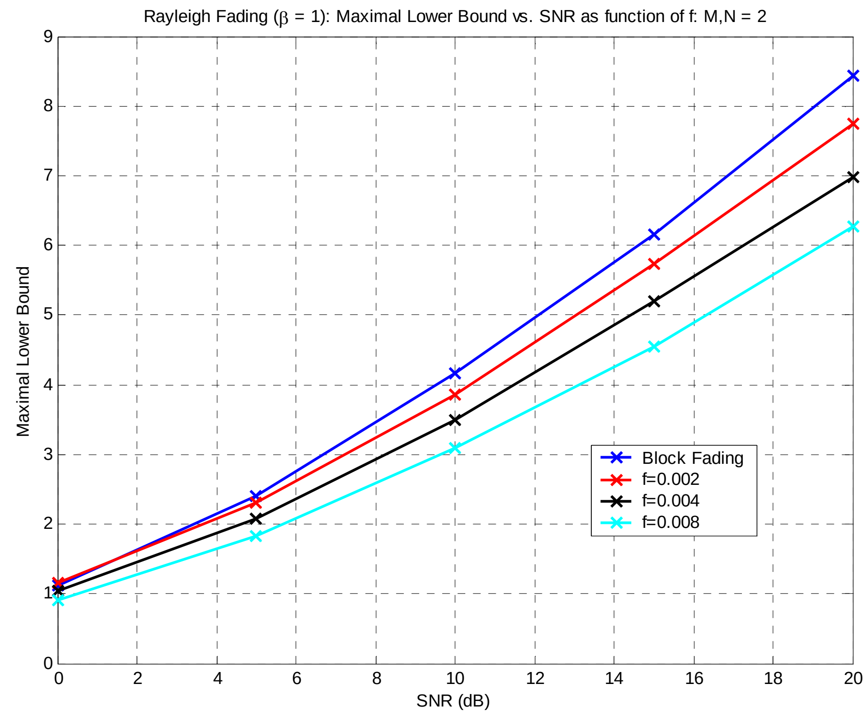
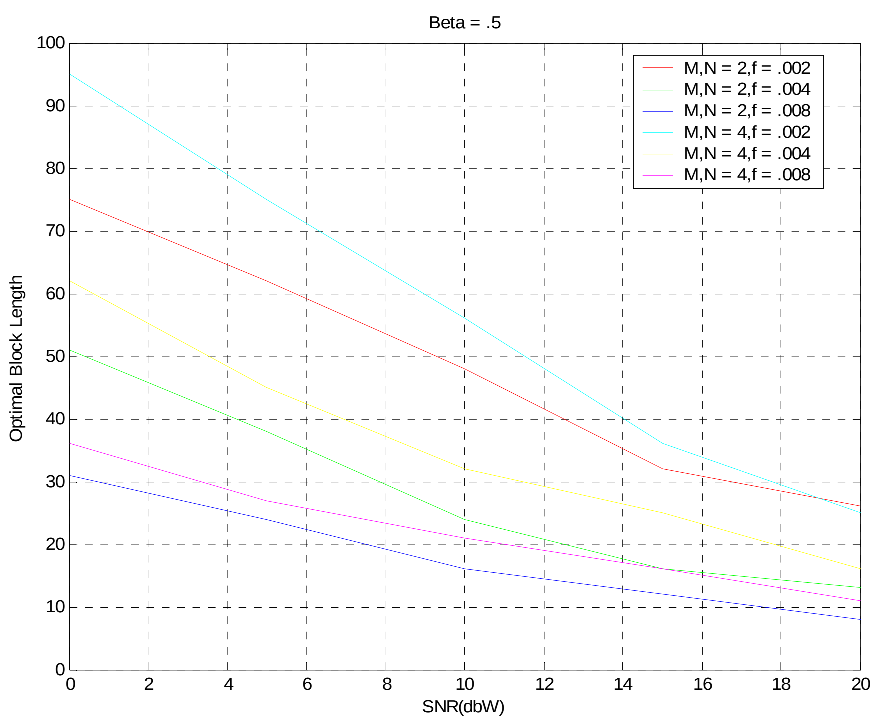
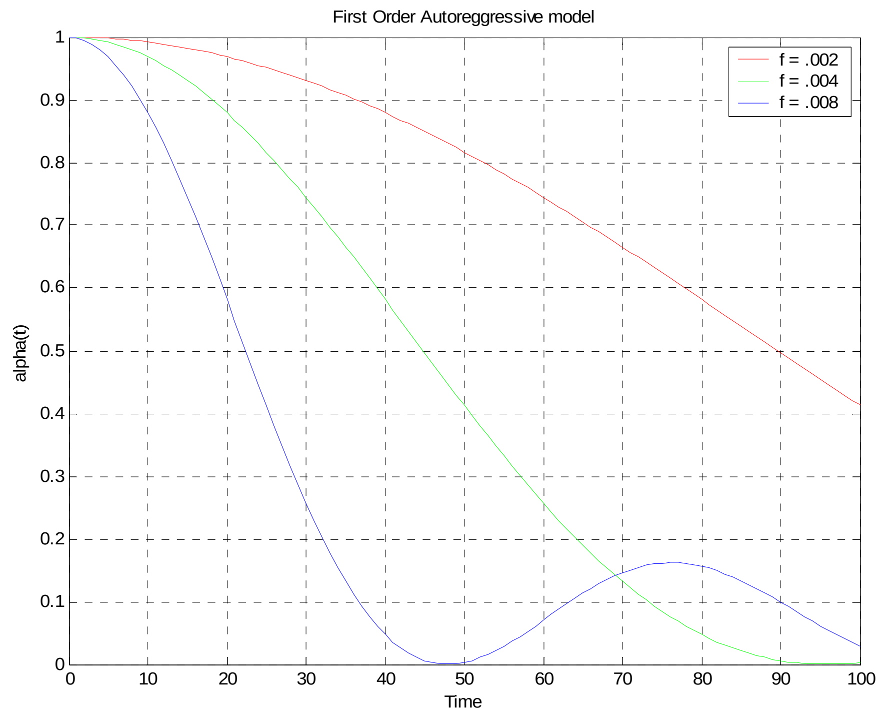


Figure 2 – Maximal Block Length vs. SNR as a function of 'f, M , and N': β = .5



<u>Figure 3 – $\alpha(t)$ vs. time as function of ‘f’:</u>



<u>Figure 4 – Optimal Training length vs. SNR as a function of ‘f, M, and N’: M,N = 2</u>

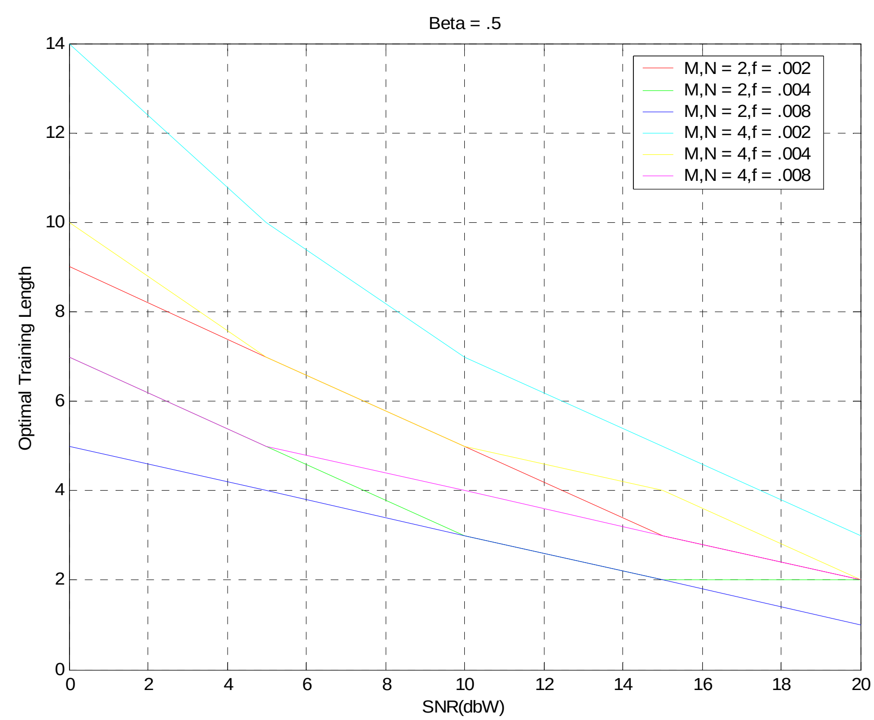


Figure 5 – Maximum Data Rate vs. SNR as a function of ‘β, M, and N’: *f* = .002