Modeling Channel Estimation Error in Continuously Varying MIMO Channels

Chris Potter (Student) and Kurt Kosbar (Advisor)
Telemetry Learning Center
University of Missouri - Rolla
Rolla, MO 65401

ABSTRACT

The accuracy of channel estimation plays a crucial role in the demodulation of data symbols sent across an unknown wireless medium. In this work a new analytical expression for the channel estimation error of a multiple input multiple output (MIMO) system is obtained when the wireless medium is continuously changing in the temporal domain. Numerical examples are provided to illustrate our findings.

KEY WORDS

Multiple-input multiple-output(MIMO), Training, Mean squared error, Payload, Temporal

INTRODUCTION

Multiple-input multiple-output (MIMO) wireless communication systems have the capability of providing a substantial increase in data rate [1, 2]. For optimal performance, one must assume that the receiver has perfect channel state information (CSI). When this is not the case, incorrect decoding at the receiver may result.

This situation motivates the use of pilot symbols to estimate the CSI. One can measure how close this estimate is to the actual channel coefficients by measuring the mean squared error (MSE) between the true channel and its estimate. In [3] the MSE was found for the Rayleigh block fading scenario, where the channel is assumed to remain constant for a block of transmitted symbols. This was extended to an arbitrary channel distribution in [4]. These works did not take into account different channel models or consider a channel that is time varying during the payload. Time variation was considered in [5], but an arbitrary channel distribution for the diffuse component was not, nor was an expression for the MSE. Moreover, the ML estimate that was used is classical in that it assumes the channel coefficients are deterministic but unknown parameters, making this method undesirable when the channel coefficients are considered stochastic [6].

In this work, a new expression for the MSE is derived for an arbitrary channel distribution that is continuously varying in time. From this result, optimal values for the amount of time to spend training are found. The optimal training time is shown to vary according to the temporal autocorrelation between channel coefficients. The rate of change of the MSE with various received SNR and number of transmit and receive antennas is investigated. Knowing how these parameters affect the MSE can be beneficial to the transmitter in a smart antenna environment or when the receiver is utilizing a low rate feedback loop to the transmitter. Particular cases of this expression are shown to yield results obtained in the literature. The case of Rayleigh fading is presented to show how the channel estimation error increases dramatically as time between training increases.

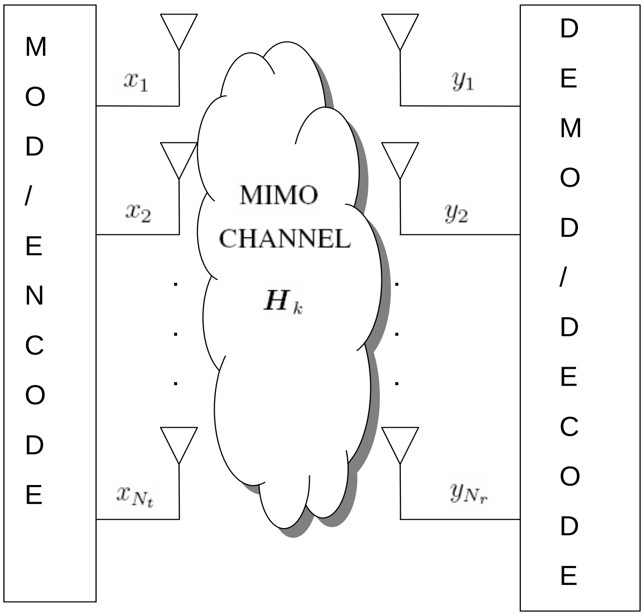

Figure 1: Block diagram for a MIMO system.

The rest of this work is presented as follows. The next section describes the mathematical models that are used to derive the main result. This is followed by previous results that relate to our work. A new expression for the channel estimation error for a continuously changing MIMO channel is derived. Optimal values for the amount of time to spend training are found and the number of antennas and received SNR are varied to show how the rate of change of the MSE is affected. Special cases of our result are shown to yield expressions that have occurred in the literature.

MATHEMATICAL MODELS

- Received symbols

Let the received symbols of a baseband MIMO channel with N_t transmitters and N_r receivers be described by

$$\boldsymbol{y}_k = \sqrt{\frac{\rho}{N_t}} \boldsymbol{H}_k \, \boldsymbol{x}_k + \boldsymbol{n}_k \tag{1}$$

where ρ is the received SNR, $\boldsymbol{y}_k$ is the N_r x 1 received symbol vector, $\boldsymbol{x}_k$ is the N_t x 1 transmitted symbol vector, and $\boldsymbol{n}_k$ is the N_r x 1 noise vector with elements $[\boldsymbol{n}]_j \overset{iid}{\sim} \mathcal{CN}(0, 1)$. The N_r x N_t time varying MIMO channel matrix $\boldsymbol{H}_k$ has indices determined by $[\boldsymbol{H}_k]_{mn} = h_{mn}(kT_s)$, where $h_{mn}(kT_s)$ is the impulse response between the m^{th} receiver and n^{th} transmitter and T_s is the symbol period. A block diagram of the overall MIMO system is illustrated in Figure 1. Throughout this work assume that the antennas are spaced enough at the transmitter and receiver so that $\boldsymbol{H}_k$ is spatially uncorrelated and is normalized such that

$$\frac{\mathbb{E}||\boldsymbol{H}_k||_F^2}{N_r N_r} = 1 \tag{2}$$

where $|| \cdot ||_F$ is the Frobenius norm [7].

- Training Phase

Suppose that each transmitter sends T_τ consecutive pilot symbols at the beginning of a T symbol block. As seen in Figure 2, the received symbols are obtained from (1) as $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_{T_\tau}$. The receiver uses a buffer to obtain $\boldsymbol{Y}_\tau = [\boldsymbol{y_1} \cdots \boldsymbol{y_{T_\tau}}]$, a N_r x T_τ matrix and also constructs the training matrix $\boldsymbol{X}_\tau = [\boldsymbol{x_1} \cdots \boldsymbol{x_{T_\tau}}]$, of size N_t x T_τ. To making training feasible, assume that the channel does not change throughout the training length and is denoted by $\boldsymbol{H}_\tau$. The received symbols during the training period can be written as

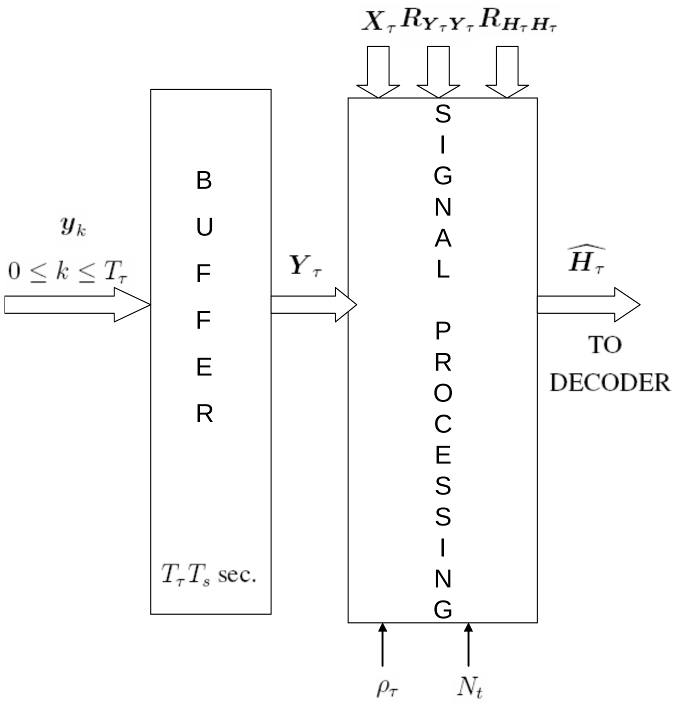

Figure 2: Block diagram of training scheme.

$$\boldsymbol{Y}_\tau = \sqrt{\frac{\rho_\tau}{N_t}} \boldsymbol{H}_\tau \boldsymbol{X}_\tau + \boldsymbol{N}_\tau \tag{3}$$

where ρ_τ is the received SNR during training and $\boldsymbol{N}_\tau$ is the noise matrix with indices $[\boldsymbol{N}_\tau]_{mn} \overset{iid}{\sim} \mathcal{CN}(0,1)$. The training matrix is restricted such that

$$||\boldsymbol{X}_\tau||_F \le T_\tau N_t. \tag{4}$$

- Payload Phase

During the payload phase the remaining $D \triangleq T - T_\tau$ symbols and $\boldsymbol{H}_k$ are not known to the receiver. In this context the received model becomes

$$\boldsymbol{Y}_t = \sqrt{\frac{\rho_d}{N_t}} \boldsymbol{H}_t \boldsymbol{X}_t + \boldsymbol{N}_t \tag{5}$$

where ρ_d is the received SNR during the payload, $\boldsymbol{Y}_t$ is of size N_r x D, $\boldsymbol{X}_t$ is of size N_t x D, and $\boldsymbol{N}_t$ is of size N_r x D with the same distribution as the training phase. The total received SNR can be written as

$$\rho = \frac{\rho_\tau T_\tau + \rho_d D}{T} \tag{6}$$

and the transmit matrix is constrained so that

$$\mathbb{E}||\boldsymbol{X}_t|| \leq N_t D. \tag{7}$$

- Channel Variation

During the data phase assume the channel varies k symbols in advance from $\boldsymbol{H}_\tau$ by [5]

$$\boldsymbol{H}_{\tau+k} = \sqrt{\alpha_k}\boldsymbol{H}_\tau + \sqrt{1-\alpha_k}\boldsymbol{W}_{\tau+k}. \tag{8}$$

where α_k is a deterministic but unknown parameter and $\boldsymbol{W}_j$ is a N_r x N_t matrix with $[\boldsymbol{W}_j]_{mn} \overset{iid}{\sim} \mathcal{CN}(0,1)$. Assuming that $\boldsymbol{W}_j$, $\boldsymbol{N}_\tau$, and $\boldsymbol{H}_\tau$ are all uncorrelated along with recognizing (8) as $N_r N_t$ independent k^{th} order autoregressive processes, we can solve for α_k as

$$\alpha_k = \left[\frac{r_{hh}(kT_s)}{r_{hh}(0)}\right]^2 \tag{9}$$

where $r_{hh}(kT_s)$ is the autocorrelation function of the channel at time kT_s.

PREVIOUS RESULTS

In [3] the authors assumed *Rayleigh block fading*, where $\boldsymbol{H}_t = \boldsymbol{H}_\tau$ for the T symbol block and $[\boldsymbol{H}_\tau]_{mn} \overset{iid}{\sim} \mathcal{CN}(0,1)$. The MSE between the $\boldsymbol{H}_\tau$ and the LMMSE estimate $\widehat{\boldsymbol{H}}_\tau$ was shown to be

$$\sigma_{bf}^2 = \frac{1}{1 + \frac{\rho}{N_t}T_\tau} \tag{10}$$

where

$$\sigma_{bf}^2 \triangleq E\{\|\boldsymbol{H}_\tau - \widehat{\boldsymbol{H}}_\tau\|_F^2\}. \tag{11}$$

In [4] the MSE for the linear minimum mean squared error (LMMSE) estimate was derived for an arbitrary channel distribution under block fading and was reported as

$$\sigma_{bf}^2 = \frac{1}{\frac{\rho T_\tau}{N_t} + \frac{N_r}{N_t}\text{trace}(\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1})}. \tag{12}$$

These results do not take into account a general channel distribution that is time varying during the payload. The contribution of this work fulfills that task.

CHANNEL ESTIMATION ERROR OF CONTINUOUSLY FADING MIMO CHANNEL

We now obtain an expression for the MSE of an arbitrary channel that is continuously varying during the payload. The MSE of the estimate is

$$\sigma_{cf}^2(k) \triangleq \mathbb{E}\{\|\boldsymbol{H}_{\tau+k} - \widehat{\boldsymbol{H}}_\tau\|_F^2\} \tag{13}$$

where $\widehat{\boldsymbol{H}}_\tau$ is the LMMSE expressed by

$$\widehat{\boldsymbol{H}}_\tau = \sqrt{\frac{\rho_\tau}{N_t}}\boldsymbol{Y}_\tau \boldsymbol{R}_{\boldsymbol{Y}_\tau \boldsymbol{Y}_\tau}\boldsymbol{X}_\tau^H \boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}. \tag{14}$$

Expanding (13), we have

$$\begin{aligned}\sigma_{cf}^2(k) &= \text{trace}(\alpha_k \boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}) \quad (15)\\ &+ \text{trace}((1-\alpha_k)N_r \boldsymbol{I}_{N_t}) + \text{trace}(\frac{\rho_\tau}{N_t}(1-2\sqrt{\alpha_k})\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}\boldsymbol{X}_\tau \boldsymbol{R}_{\boldsymbol{Y}_\tau \boldsymbol{Y}_\tau}^{-1}\boldsymbol{X}_\tau^H \boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau})\\ &= (\alpha_k + 1 - 2\sqrt{\alpha_k})\text{trace}(\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}) + (1-\alpha_k)N_r N_t\\ &- (1-2\sqrt{\alpha_k})\text{trace}([\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1} + \frac{\rho_\tau}{N_t N_t}\boldsymbol{X}_\tau \boldsymbol{X}_\tau^H]^{-1}). \quad (16)\end{aligned}$$

We now find find a training matrix such that $\boldsymbol{X}_\tau \boldsymbol{X}_\tau^H$ is non-singular which minimizes (16). Since only the third term depends on $\boldsymbol{X}_\tau$, the minimization can be stated as

$$\arg\min_{\boldsymbol{X}_\tau^H} \text{trace}([\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1} + \frac{\rho_\tau}{N_t N_t}\boldsymbol{X}_\tau \boldsymbol{X}_\tau^H]^{-1}) \quad (17)$$

such that

$$\text{trace}(\boldsymbol{X}_\tau \boldsymbol{X}_\tau^H) = N_t T_\tau. \quad (18)$$

Setting the derivative of (27) with respect to $\boldsymbol{X}_\tau^H$ equal to zero it can be shown that

$$\boldsymbol{X}_\tau \boldsymbol{X}_\tau^H = \left(\frac{N_r}{\rho_\tau}\text{trace}(\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1}) + T_\tau\right)\boldsymbol{I}_{N_t} - \frac{N_t N_r}{\rho_\tau}\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1}. \quad (19)$$

Plugging (19) into (16), simplifying, and taking (2) into account the normalized MSE can be expressed by

$$\widetilde{\sigma_{cf}}^2(k) \triangleq \frac{\sigma_{cf}^2(k)}{N_r N_t} = 2(1-\sqrt{\alpha_k}) - \frac{(1-2\sqrt{\alpha_k})}{\frac{\rho_\tau T_\tau}{N_t} + \frac{N_r}{N_t}\text{trace}(\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1})}. \quad (20)$$

Note that the noise term used to model the channel variation in (8) results in $0 \leq \widetilde{\sigma_{cf}}^2(k) \leq 2$.

- Optimal Training Length

The optimal T_τ when N_t, N_r, and ρ_τ, are fixed is attained. The partial derivative of (20) with respect to T_τ can be expressed as

$$\begin{aligned}\frac{\partial \widetilde{\sigma_{cf}}^2(k)}{\partial T_\tau} &= -\frac{\partial}{\partial T_\tau}\left[\frac{(1-2\sqrt{(\alpha_k)})}{\frac{\rho_\tau T_\tau}{N_t} + \frac{N_r}{N_t}\text{trace}\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1}}\right]\\ &= \frac{(1-2\sqrt{\alpha_k})}{(\frac{T_\tau N_t}{\rho_\tau} + \frac{N_r N_t}{\rho_\tau^2}\text{trace}\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}^{-1})^2}. \quad (21)\end{aligned}$$

Observe that since the denominator in (21) is positive, the sign of the derivative is completely dependent on α_k. Consider the following cases.

1. $\alpha_k = 1/4$: Clearly $\frac{\partial \widetilde{\sigma_{cf}}^2(k)}{\partial T_\tau} = 0$ which means $\widetilde{\sigma_{cf}}^2(k)$ is constant for every T_τ, which suggests choosing $T_\tau = 0$. The channel in this case has as much noise as autocorrelation, making an accurate estimate unpractical.

2. $\alpha_k < 1/4$: Since the derivative is positive $\widetilde{\sigma_{cf}}^2(k)$ is monotonically increasing, inferring the choice of $T_\tau = 0$. This is justified by noting that when α_k is decreased, the time varying channel becomes more erratic from sample to sample making the training estimate useless.

3. $\alpha_k > 1/4$: Through similar reasoning as Case 2, $\widetilde{\sigma_{cf}}^2(k)$ is monotonically decreasing which justifies making T_τ as large as possible. When the channel coefficients are highly autocorrelated in the temporal sense, the channel is less susceptible to variation, making it beneficial to spend sufficient time learning the channel

We now investigate how the channel estimation error is affected for various N_r, N_t and ρ_τ. Looking at (21), it is clear that increasing N_r or N_t would result in decreasing the rate of change of the MSE, whereas increasing ρ_τ will have the opposite effect.

To illustrate these results with a simple example (20) and (21) have been tabulated for various values of ρ_τ, α_k, and N_t under the assumption of iid Rayleigh fading. Looking at Figures 3 and 4 one can see the graphs validate the optimal training values for the different values of α_k and also exemplify how the rate of change of the MSE varies with ρ_τ, N_t, and N_r.

We now summarize these results into the following proposition.

Proposition 1 *Given a MIMO continuously varying channel with parameters ($\alpha_k, \rho_\tau, N_r, N_t, T_\tau$) the optimal time to spend training is*

1. $T_\tau = 0$ *when* $\alpha_k \leq 1/4$.
2. $T_\tau = D$ *when* $\alpha_k > 1/4$.

Furthermore increasing N_t and N_r reduces the rate of change of the MSE, whereas increasing ρ_τ accelerates the rate of change.

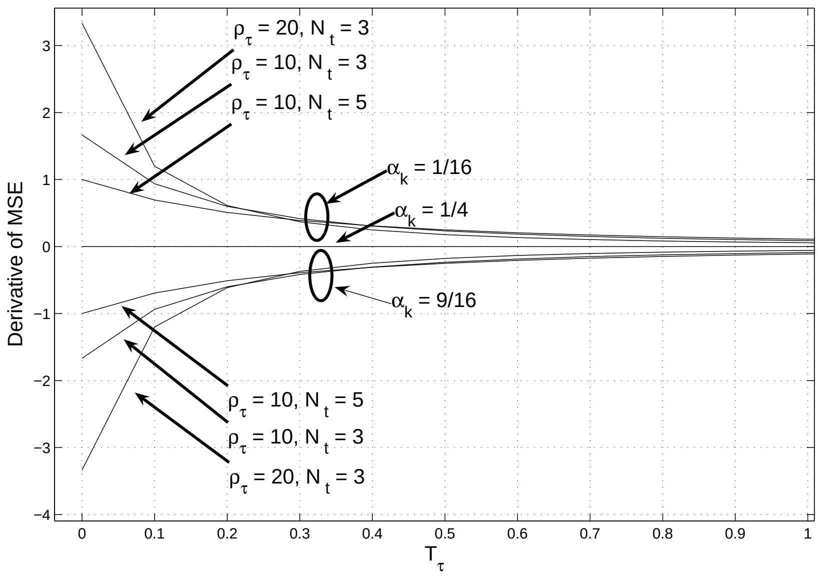

Figure 3: Variation in the partial derivative with respect to T_τ for various α_k, N_t, and ρ_τ.

SPECIAL CASES

In this section special cases of our new result for the channel estimation error for a continuously varying MIMO channel are presented which yield previous forms in the literature. These are followed by the uncorrelated block fading and the Rayleigh continuously varying cases.

- **Rayleigh Block Fading**

Let $\alpha_k = 1$ for all k and $\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau} = N_r \boldsymbol{I}_{N_t}$. Then (20) reduces to

$$\widetilde{\sigma_{cf}}^2(k) = \frac{1}{1 + \frac{\rho}{N_t} T_\tau}\,, \quad k = T_\tau + 1, \ldots, T \tag{22}$$

which is precisely (10).

- **Arbitrary Block Fading**

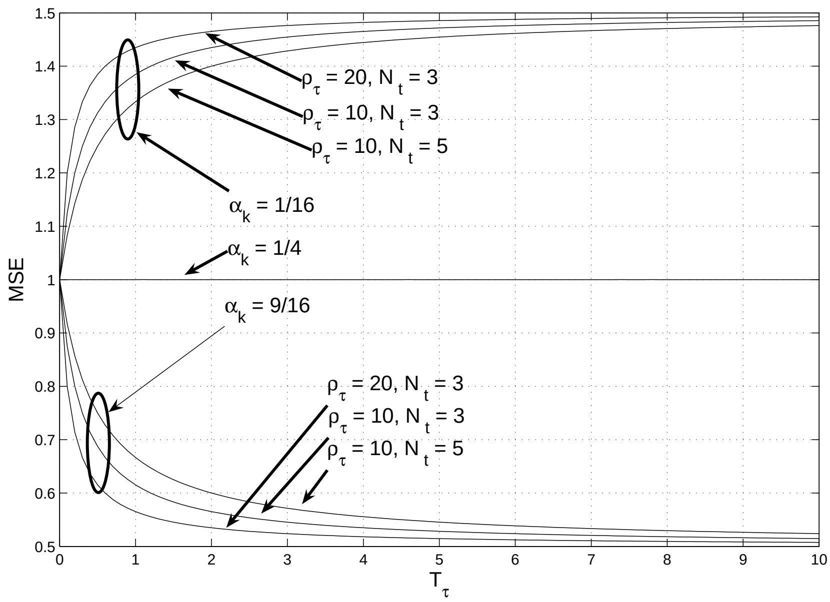

Figure 4: Variation in MSE with respect to T_τ for various α_k, N_t, and ρ_τ.

Suppose $\alpha_k = 1$ for all k. Then (20) becomes

$$\widetilde{\sigma_{cf}}^2(k) = \frac{1}{\frac{\rho T_\tau}{N_t} + \frac{N_r}{N_t}\mathrm{trace}(\boldsymbol{R}^{-1}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau})}\,, \quad k = T_\tau + 1, \ldots, T \tag{23}$$

which is exactly (11).

- **Uncorrelated Block Fading**

Let $\alpha_k = 0$ for all k. Then (20) can be written as

$$\widetilde{\sigma_{cf}}^2(k) = 1 + \frac{\mathrm{trace}(\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau})}{N_r N_t} - \frac{1}{\frac{\rho_\tau T_\tau}{N_t} + \frac{N_r}{N_t}\mathrm{trace}(\boldsymbol{R}^{-1}_{HH})}\,. \tag{24}$$

We now seek the $\boldsymbol{H}_\tau$ that minimizes (24). It can be shown [6] that

$$\mathrm{trace}(\boldsymbol{R}^{-1}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}) \geq \sum_i \frac{1}{[\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}]_{ii}} \tag{25}$$

with equality if and only if $\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}$ is diagonal. We now seek to find the optimal diagonal matrix that minimizes

$$\arg \min_{\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}} \mathrm{trace}(\boldsymbol{R}^{-1}_{HH}) \tag{26}$$

such that

$$\mathbb{E}||\boldsymbol{H}_\tau||_F = N_r N_t \ . \tag{27}$$

Differentiating with respect to $\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau}$, setting the result equal to zero and applying the constraint we obtain

$$\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau} = N_r \boldsymbol{I}_{N_t} \tag{28}$$

which shows that the MSE is minimized under time-uncorrelated block fading when the channel coefficients are uncorrelated.

- Rayleigh Continuously Varying

Now let $\boldsymbol{R}_{\boldsymbol{H}_\tau \boldsymbol{H}_\tau} = N_r \boldsymbol{I}_{N_t}$. The channel estimation error can be written as

$$\widetilde{\sigma_{cf}}^2(k) = \frac{2(1-\sqrt{\alpha_k})}{1+\frac{N_t}{\rho_\tau T_\tau}} + \frac{1}{\frac{\rho_\tau T_\tau}{N_t}+1} \tag{29}$$

The MSE is expressed as the sum of two terms, the first is the error due to the time varying nature of the channel and the second is the Rayleigh block fading MSE previously mentioned. The Rayleigh block fading component can be thought of the error due to channel estimation during the training period. Unlike the block fading term, an increase in ρ_τ or T_τ adversely affects the time varying term, necessitating a tradeoff between the two components.

In order to demonstrate the difference, the MSE for the Rayleigh block fading case and the Rayleigh continuous varying case is tabulated. Assume that the temporal variation in the channel obeys Jake's model resulting in

$$\alpha_k = J_o(2\pi f_d T_s)^2 \tag{30}$$

where f_d is the doppler frequency and $J_o(\cdot)$ is the zeroth order bessel function of the first kind. Let $T_s = 1\mu$ sec, $f_d = 200$ Hz, and $N_r = N_r = 3$. Looking at Figure 5 one can see that the continuously fading MSE does not deviate substantially from the block fading case. This is because the normalized frequency $f_d T_s$ is small, resulting in a high value for α_k during the payload. This keeps the time-varying component from fluctuating.

Suppose now that $T_s = 1$msec. The results in Figure 6 indicate that the continuously fading MSE varies drastically from the block fading case. The normalized frequency is now three orders of magnitude larger than the previous case, causing α_k to fluctuate during the payload. This in turn causes the time varying component to fluctuate more rapidly, leading to an increase in the MSE.

CONCLUSION

Previous expressions for the channel estimation error did not take into account an arbitrary channel that varies temporally during the payload. In this work an expression for the channel estimation error under this condition was derived. Optimal values for the training time were found and it was shown how varying the received SNR and the number of antennas affected the rate of change of the channel estimation error. Special cases of our new result were shown to result in previous results in the literature. For Rayleigh fading it was shown that varying the normalized frequency impacted the accuracy of the channel estimate and hence the channel estimation error.

REFERENCES

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

[2] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Syst. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[3] 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.

[4] M. Biguesh and A. Gershman, "Training-based mimo channel estimation: A study of estimator tradeoffs and optimal training signals." *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 1053–1058, Mar. 2006.

[5] 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.

[6] S. Kay, *Fundamentals of Statistical Signal Processing:Estimation Theory*. Upper Saddle River,NJ: Prentice Hall, 1993.

[7] J. M. Ortega, *Matrix Theory-A Second Course*. New York, NY: Plenum, 1987.

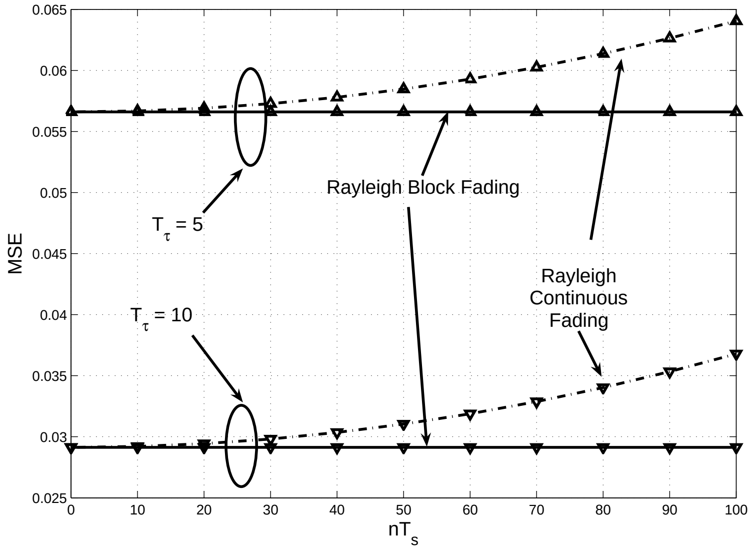

Figure 5: Channel estimation error comparison for the Rayleigh block fading case and the Rayleigh continuously fading case as a function of symbol period for various T_τ when $N_r = N_t = 3$,$f_d = 200$ Hz, and $T_s = 1\mu$ sec.

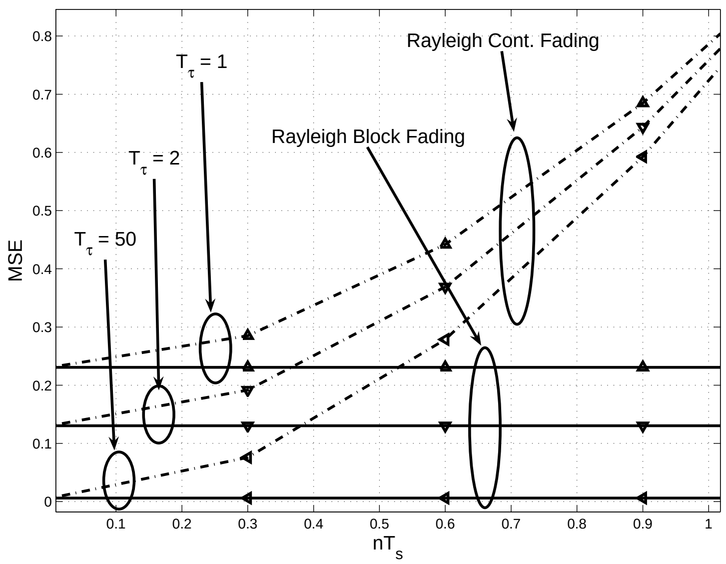

Figure 6: Channel estimation error comparison for the Rayleigh block fading case and the Rayleigh continuously fading case as a function of symbol period for various T_τ when $N_r = N_t = 3$,$f_d = 200$ Hz, and $T_s = 1$ msec.