

# Optimal Trained Space-Time Modulation over a Rician Time-Varying Channel

Christian B. Peel and A. Lee Swindlehurst  
Brigham Young University  
Electrical and Computer Engineering Dept.  
459 CB, Provo, UT 84602

chris.peel@ieee.org, swindle@ee.byu.edu

**Abstract**—The power allocation, training signal length, and training signal which maximize a lower bound on the capacity of a Rician time-varying channel are found in this paper. We consider the case where the channel consists of two components: a specular component from line-of-sight or strong rays, and a time-varying diffuse component due to a mobile receiver or scatterers. Trained modulation is the most common technique for accommodating such time-varying channels. We optimize a lower bound on the capacity to find the parameters mentioned above. Our results are illustrated with several numerical examples.

## I. INTRODUCTION

Interest in time-varying multiple-antenna channels has increased recently, motivated by demand for high bandwidth mobile wireless data links. This has prompted investigation of modulation techniques that can handle temporally selective fading. Though trained modulation is not ideal for a time-varying channel, it is often used because of its simplicity. The parameters that optimize a lower bound on capacity for a training-based scheme have been found by Hassibi and Hochwald [1]. They assume quasi-static fading, or that the channel coefficients are constant for a given coherence interval before changing to new independent values. The optimal training signal is found to be unitary, and the optimal training signal length, and training power are also found.

Another approach for time-varying fading is to assume that the receiver has no knowledge of the channel. This is in contrast to receivers for trained modulation, for example, which assume that the channel is known by the receiver. In the case of no channel knowledge, unitary signals have been found to be optimal [2]. Several signaling schemes have been developed using this assumption, including differential space-time modulation [3], [4]. The capacity in the high-SNR case has been found [5] which provides insight on the work in [1]. Related work in [6] gives the length of training signal needed for a frequency selective channel.

A different approach was taken in [9] where the time-autocorrelation function of the channel coefficients, and mul-

multiple first-order AR systems are used to model the channel. An effective SNR that includes the penalty due to channel estimation and time variation is found. In this paper, we use this channel model to derive a lower bound on capacity for training-based systems. This bound is the extension of results in [1] for a time-varying Rician channel with known autocorrelation function. We find the training signal length, training power, and training signal that maximize the bound. The next section presents the channel model, including a description of channel estimation techniques used. In Section III we derive the lower bound; we conclude with plots of the capacity as a function of the optimized parameters.

## II. CHANNEL MODEL

In what follows, we let  $\mathcal{CN}(0, 1)$  denote a zero-mean, unit-variance, circularly symmetric complex Gaussian distribution. The  $T \times T$  identity matrix is indicated by  $I_T$ , the Frobenius norm by  $\|\cdot\|_F$ , the expectation operator by  $E[\cdot]$ , and the determinant by  $|\cdot|$ . We say an  $M \times N$  matrix  $A$  has unit mean-square if

$$E[\|A\|_F^2] = E[\text{tr}(AA^H)] = MN.$$

### A. Fading Channel Model

Assume a flat-fading communications environment with  $M$  transmit and  $N$  receive antennas. A complex channel coefficient that is assumed constant for a symbol of length  $T_s$  describes the effect of the propagation between each pair of transmit and receive antennas. These channel coefficients are assumed to be independent from element to element across the antenna array. At each receive antenna, interference and other disturbances add temporally and spatially independent noise to the signal. In this case, the channel equation can be written in matrix form:

$$X = \sqrt{\frac{p}{M}}SH + V, \quad (1)$$

where  $X$  is a  $T_s \times N$  matrix of received data,  $S$  is a  $T_s \times M$  signal matrix,  $H$  is a  $M \times N$  matrix of channel coefficients,

and  $V$  is a  $T_s \times N$  matrix of additive receiver noise. The matrices  $S$ ,  $H$ , and  $V$  all are assumed to have unit mean-square; the expected received SNR at each receive antenna is thus  $\rho$ . Though we will derive several initial results that do not depend on the distribution of  $H$ , the majority of this work will assume that  $H$  has i.i.d.  $\mathcal{CN}(0, 1)$  entries.

### B. Channel Estimation

We apply this fading model to both training and data transmission phases, resulting in the following joint channel model

$$\begin{bmatrix} X_\tau \\ X_d \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{\rho_\tau}{M}} S_\tau H_\tau \\ \sqrt{\frac{\rho_d}{M}} S_d H_d \end{bmatrix} + \begin{bmatrix} V_\tau \\ V_d \end{bmatrix}, \quad (2)$$

where the training signal  $S_\tau$  has dimension  $T_\tau \times M$  and the data signal has dimension  $T_d \times M$ . Conservation of time and energy yields the following:

$$T = T_\tau + T_d \quad (3a)$$

$$\rho T = \rho_\tau T_\tau + \rho_d T_d, \quad (3b)$$

where  $\rho_\tau$  and  $\rho_d$  are the relative signal powers employed during training and data transmission, respectively. An important issue that we will address is the allocation of energy between training and data transmission. We use the parameter  $\kappa$  to make the tradeoff:

$$\rho_d T_d = \kappa \rho T \quad (4a)$$

$$\rho_\tau T_\tau = (1 - \kappa) \rho T. \quad (4b)$$

The penalty due to channel estimation is also a major focus of this work. We consider the MMSE estimator:

$$\hat{H}_{\text{MMSE}} = \sqrt{\frac{\rho_\tau \beta}{M}} \left( I_M + \frac{\rho_\tau \beta}{M} S_\tau^H S_\tau \right)^{-1} S_\tau^H \begin{pmatrix} X_\tau - \sqrt{(1 - \beta) \frac{\rho_\tau}{M}} S_\tau H^s \end{pmatrix},$$

### C. Specular and Diffuse Components

The channel will be split into specular and diffuse components as follows:

$$H = \sqrt{1 - \beta} H^s + \sqrt{\beta} H^r, \quad (5)$$

where the specular part  $H^s$  is assumed to be time-invariant and known, while the diffuse component  $H^r$  is zero-mean and time-varying. The specular and diffuse power are also often separated as follows

$$\rho^s = (1 - \beta) \rho, \quad \rho^r = \beta \rho, \quad (6)$$

where  $\rho = \rho^s + \rho^d$ . The parameter  $\beta$  allows tuning between a fully specular ( $\beta = 0$ ) and diffuse ( $\beta = 1$ ) channel. In practice we expect the channel to be composed of both diffuse and specular components, in which case  $0 < \beta < 1$ . For most of the paper we specify no specific distribution on the channel except that it be left and right rotationally

invariant, i.e., when  $p(\Theta H \Phi) = p(H)$  for all  $\Theta$  and  $\Phi$  such that  $\Theta^H \Theta = \Theta \Theta^H = I_M$  and  $\Phi^H \Phi = \Phi \Phi^H = I_N$ .

As an example of a channel that satisfies this model, we now describe a Rician channel with rank-one specular component; in this case  $H^s$  is written as the outer product of two known unit vectors:

$$H^s = \sqrt{MN} \mathbf{u} \mathbf{v}^H, \quad (7)$$

where  $\mathbf{u}^H \mathbf{u} = 1$  and  $\mathbf{v}^H \mathbf{v} = 1$ . The diffuse component in this case is composed of  $\mathcal{CN}(0, 1)$  random variables. Rank-one models similar to this have been used for analysis of capacity in [8], and analysis of time-varying channels in [9]; this model will be assumed for the numerical examples concluding the paper.

### D. Time-Varying Diffuse Element

The quasi-static model of a time-varying channel assumes that  $T$  is the coherence length of the channel, or that the channel is approximately constant for  $T$  samples. In our model we go one step further, by allowing the channel to change between the data and training phases of transmission. This means that  $T$  is no longer a coherence interval, but is still the coding interval or training period. We assume that the channel varies according to the following first-order autoregressive (AR) or *Gauss-innovations* model:

$$H_d^r = \sqrt{\alpha_T} H_\tau^r + \sqrt{1 - \alpha_T} W_d, \quad (8)$$

where  $H_\tau^r$  and  $W_d$  have the same distribution, are zero-mean, and have unit mean-square; we also constrain  $0 \leq \alpha_T \leq 1$ . Under this model,  $H_d$  has the same distribution as  $H_\tau$ . Note that  $\alpha_T = 1$  produces a time-invariant channel, and  $\alpha_T = 0$  indicates a channel with no temporal correlation.

The parameter  $\alpha_T$  can be chosen to match the second order statistics of models based on the mechanisms of physical propagation. Let  $r_{hh}(t)$  denote the time autocorrelation function of an element of  $H_d^r$ . Solving the Yule-Walker equations for  $\alpha_T$  in the first-order AR process (8) we obtain

$$\alpha_T = \left[ \frac{r_{hh}(T/2)}{r_{hh}(0)} \right]^2, \quad (9)$$

which provides a reasonable choice for  $\alpha_T$ . For example, assuming Jakes' model of the land mobile fading channel [10],  $r_{hh}(t) = J_0(2\pi f t)$ , where  $J_0(\cdot)$  is the zeroth-order Bessel function of the first kind,  $f = f_d T_s$ ,  $f_d$  is the maximum Doppler frequency in the fading environment, and  $T_s$  is the sampling period. Under this model (9) leads to

$$\alpha_T = J_0(\pi T f)^2. \quad (10)$$

The Gauss-innovations model is an appropriate approximation when using decoders that depend only on a single reference channel (i.e., decoders that assume the channel is known). This fact is borne out by the simulation results of [9], where excellent agreement is obtained with data generated according to Jakes' model, but analyzed with the Gauss-innovations model.

### III. CAPACITY

An analysis of the capacity achievable with trained modulation over the time-varying channel described above is now presented. A lower bound on capacity is derived, extending the work of [1] to a time-varying channel with known autocorrelation function. We find that while the lower bound is significantly different than that in [1], the parameters that optimize the bound are not.

#### A. Capacity and Effective SNR

Section II describes in detail the relationship between the channel seen by the received and training data:

$$\begin{aligned} X_d &= \sqrt{\frac{\rho_d}{M}} S_d H_d + V_d \\ &= \sqrt{(1-\beta) \frac{\rho_d}{M}} S_d H^s + \sqrt{\beta \alpha_T \frac{\rho_d}{M}} S_d \hat{H}_d^r + \\ &\quad \sqrt{\beta(1-\alpha_T) \frac{\rho_d}{M}} S_d W_d + V_d. \end{aligned}$$

The difference between the estimated diffuse channel and that seen by the training data is  $\hat{H}_d^r = H_d^r - \hat{H}_d^r$ . We can now separate the received signal from the effective noise:

$$\begin{aligned} X_d &= \underbrace{\sqrt{(1-\beta) \frac{\rho_d}{M}} S_d H^s + \sqrt{\beta \alpha_T \frac{\rho_d}{M}} S_d \hat{H}_d^r}_{X'_d} + \underbrace{\sqrt{\beta \alpha_T \frac{\rho_d}{M}} S_d \hat{H}_d^r + \sqrt{\beta(1-\alpha_T) \frac{\rho_d}{M}} S_d W_d + V_d}_{V'_d}. \end{aligned} \quad (11)$$

The variance of the effective noise term is found to depend on the variance of the channel estimation error:

$$\sigma_{V'_d}^2 = \frac{1}{T_d N} E \text{tr} \left[ V_d'^H V_d' \right] = \rho_d \alpha \beta \sigma_{\hat{H}_d^r}^2 + \rho_d \beta (1-\alpha) + 1. \quad (12)$$

We also note that the variance of  $X'_d$  is

$$\sigma_{X'_d}^2 = \frac{1}{T_d N} E \text{tr} \left[ X_d'^H X_d' \right] = \rho_d (1-\beta) + \rho_d \alpha \beta \sigma_{\hat{H}_d^r}^2. \quad (13)$$

It is useful to compare the equations for  $X_d$  in (2) and (11). In the former the channel contains an unknown diffuse component, while in (11) the channel components  $H^s$  and  $\hat{H}_d^r$  are completely known. Theorem 1 from [1] provides a means to bound the capacity for this training based scheme. It indicates that the capacity for a training based system with known channel that is rotationally invariant is lower bounded by a system with i.i.d. zero-mean Gaussian noise. We can then write

$$C_r \leq \frac{T - T_r}{T} E \left[ \log \left| I_N + \rho_{\text{eff}} \frac{\hat{H}^H \hat{H}}{M} \right| \right], \quad (14)$$

where  $\hat{H}$  has as specular components  $H^s$  and diffuse component  $\hat{H}_d^r$ , with specular parameter

$$\beta_{\text{eff}} = \frac{\alpha \beta \sigma_{\hat{H}_d^r}^2}{1 - \beta + \alpha \beta \sigma_{\hat{H}_d^r}^2},$$

and effective SNR

$$\rho_{\text{eff}} = \frac{\rho_d (1-\beta) + \rho_d \alpha \beta \sigma_{\hat{H}_d^r}^2}{1 + \rho_d \beta (1-\alpha) + \rho_d \alpha \beta \sigma_{\hat{H}_d^r}^2}. \quad (15)$$

The remainder of this section will focus on maximizing this lower bound.

Inspection of (14) reveals that the only place that the signal enters the equation is through the effective SNR term. Thus to maximize the lower bound on capacity, we need only maximize  $\rho_{\text{eff}}$ . From the characteristics of our estimator, we know that

$$\sigma_{\hat{H}_d^r}^2 = 1 - \sigma_{H_d^r}^2.$$

This allows a different form for  $\rho_{\text{eff}}$ :

$$\rho_{\text{eff}} = \frac{\rho_d + 1}{1 + \rho_d \beta (1-\alpha) + \rho_d \alpha \beta \sigma_{H_d^r}^2} - 1. \quad (16)$$

Thus, minimizing  $\sigma_{H_d^r}^2$  also maximizes (14); this problem has already been solved in [1] where the optimal training signal is found to satisfy

$$S_r^H S_r = T_r I_M. \quad (17)$$

We will assume unitary training signals throughout the remainder of this work; in this case we write

$$\sigma_{H_d^r}^2 = \frac{\rho_r \beta \frac{T_r}{M}}{1 + \rho_r \frac{T_r}{M}} \quad (18)$$

and the effective SNR becomes

$$\rho_{\text{eff}} = \frac{\rho_d (1-\beta) + \rho_d \alpha \beta \frac{\rho_r \beta \frac{T_r}{M}}{1 + \rho_r \beta \frac{T_r}{M}}}{\rho_d \alpha \beta \frac{1}{1 + \rho_r \beta \frac{T_r}{M}} + \rho_d \beta (1-\alpha) + 1} \quad (19)$$

$$= \frac{\rho_d (1-\beta) (M + \beta \rho_r T_r) + \rho_d \rho_r \alpha \beta^2 T_r}{(M + \rho_r T_r \beta) (1 + \rho_d \beta) - \alpha \rho_d \rho_r T_r \beta^2}. \quad (20)$$

In addition, the new beta value becomes

$$\beta_{\text{eff}} = \frac{\alpha \rho_r T_r \beta^2}{(1-\beta) (M + \beta \rho_r T_r) + \alpha \rho_r T_r \beta^2}. \quad (21)$$

We now turn our attention to optimizing the power allocation. Allocating more power to training means a better channel estimate, but less power available for data transmission.

*Theorem 1 (Optimal Power Allocation):* Given the channel model of Section II and unitary training signals, the

optimal power allocation for space-time modulation is  $\kappa = \min(1, \hat{\kappa})$ , where

$$\hat{\kappa} = \begin{cases} \gamma - \sqrt{M\alpha\gamma \left[ \frac{\gamma}{T_\beta} + \frac{T_d}{\rho T T_\beta} \right]} & \text{for } T_\beta > \alpha M \\ \frac{\rho T \beta (1 - \beta + \alpha \beta) + M(1 - \beta)}{\rho T \beta (1 - \beta + \alpha \beta)} & \text{for } T_\beta = \alpha M \\ \gamma + \sqrt{M\alpha\gamma \left[ \frac{\gamma}{T_\beta} + \frac{T_d}{\rho T T_\beta} \right]} & \text{for } T_\beta < \alpha M \end{cases}, \quad (22)$$

$T_\beta = T_d(1 - \beta + \alpha\beta)$ , and

$$\gamma = \frac{T_\beta(M + \rho T \beta)}{\rho T \beta (T_\beta - \alpha M)}. \quad (23)$$

Substituting the power allocation into (19) and (21) gives the effective SNR and effective  $\beta$  values, respectively.

*Proof:* The power allocation enters into the capacity bound (14) only through the effective SNR term. Maximizing (19) with respect to  $\kappa$ , the desired result is obtained. ■

When  $\kappa = 1$ , all power is allocated to data transmission, and no power to training. In this case, the training length  $T_\tau$  is set to 0, the effective specular parameter  $\beta_{\text{eff}} = 0$  and transmission occurs only over the specular component of the channel.

At extreme SNR the power allocation and effective SNR simplify considerably: at high SNR  $\gamma = T_\beta / (T_\beta - \alpha M)$ , while at low SNR

$$\gamma = \frac{M T_\beta}{\rho T \beta (T_\beta - \alpha M)},$$

which lead to the following result.

*Corollary 1:* At high SNR, the power allocation and effective SNR become

$$\kappa = \min \left( 1, \frac{\sqrt{T_\beta}}{\sqrt{T_\beta} + \sqrt{\alpha M}} \right), \quad (24)$$

while the the effective power and specular parameter are

$$\rho_{\text{eff}} = \frac{1 - \beta + \alpha \beta}{(1 - \alpha)\beta}, \quad \beta_{\text{eff}} = \frac{\alpha \beta}{1 - \beta + \alpha \beta}. \quad (25)$$

At low SNR all power is allocated to data transmission ( $\kappa = 1$ ) and the effective SNR and specular parameter are

$$\rho_{\text{eff}} = \rho(1 - \beta), \quad \beta_{\text{eff}} = 0. \quad (26)$$

The best training interval  $T_\tau$  is not obvious. Longer intervals result in higher effective SNR values, while shorter intervals leave more time for data transmission. Note that  $T_\tau$  must be at least  $M$  in order to obtain a channel estimate.

*Theorem 2 (Optimal Training Interval):* Assume unitary training signals, optimal power allocation, and  $T_\beta = \alpha M$  or asymptotic SNR approximations from Corollary 1. The best training interval is  $T_\tau = M$ , except when  $\kappa = 0$ , in which case the best training length is  $T_\tau = 0$ <sup>1</sup>.

<sup>1</sup>The same result appears to hold for  $T_d > M$  and  $T_d < M$  at any SNR, except for when  $\beta$  is small, in which case no training should be used.

*Proof:* Let  $Q = \min(M, N)$  and  $\lambda$  be an arbitrary eigenvalue of the matrix  $\frac{\tilde{H}^H \tilde{H}}{M}$ . Then the bound (14) can be expressed as

$$C_\tau \leq \frac{Q T_d}{T} E \log_2(1 + \rho_{\text{eff}} \lambda) \triangleq C_t. \quad (27)$$

We will find the  $T_d$  that will maximize  $C_t$ , and use that result to determine the best training data length. The derivative of  $C_t$  with respect to  $T_d$  is

$$\frac{dC_t}{dT_d} = \frac{Q}{T} E \log(1 + \rho_{\text{eff}} \lambda) + \frac{Q T_d}{T} \frac{d\rho_{\text{eff}}}{dT_d} E \left[ \frac{\lambda}{1 + \rho_{\text{eff}} \lambda} \right]. \quad (28)$$

Assuming  $T_\beta = \alpha M$ , or asymptotic SNR approximations from Corollary 1 we obtain

$$\frac{d\rho_{\text{eff}}}{dT_d} = 0. \quad (29)$$

This means that

$$\frac{dC_t}{dT_d} = \frac{Q}{T} E \log(1 + \rho_{\text{eff}} \lambda) \geq 0, \quad (30)$$

and making  $T_d$  as large as possible, or  $T_\tau$  as small as possible ( $T_\tau = M$  or  $T_\tau = 0$ ), maximizes the lower bound  $C_t$ . ■

The model presented in Section II allows the determination of the best training frequency which depends on the power, number of transmit antennas, specular parameter, and autocorrelation function of the channel coefficients:

$$\hat{T} = \arg \max_T \frac{T - M}{T} E \left[ \log \left| I_N + \rho_{\text{eff}} \frac{\tilde{H}^H \tilde{H}}{M} \right| \right]. \quad (31)$$

Because of the nature of the autocorrelation function and the “log-det” capacity expression, a closed form solution for  $\hat{T}$  is difficult to obtain. Using a high-SNR approximation for the capacity, values for  $\hat{T}$  can be found in the Rayleigh case [11], but no similar approximation is currently available for the general Rician channel. We present simulation results in the next section using Monte-Carlo evaluation of (31).

#### IV. NUMERICAL EXAMPLES

Numerical results are now presented with comparisons to the analysis above. Each simulation point in the following plots is the result of a Monte-Carlo evaluation of (14) with 200 channel matrices. Figure 1 shows capacity as a function of power allocation factor  $\kappa$  for  $\rho = 20$ dB SNR,  $M = N = 2$  antennas, training interval of  $T_\tau = 2$ , normalized Doppler frequency  $f = 0.003$  and specular parameter  $\beta = 0.4$ . The maximum predicted by analysis is  $\kappa = .74$  for  $T = 18$ , which agrees well with simulation. The capacity for a system with perfect channel state information at the receiver (CSIR) is also shown.

Using the optimal power allocation,  $\rho = 20$ dB SNR,  $N = M = 2$  antennas, normalized Doppler frequency  $f = 0.003$ , training frequency  $T = 18$ , and specular parameter  $\beta = 0.4$ , the capacity as a function of the training length is shown in Figure 2. Analysis predicts that the best  $T_\tau$  to use is 2, which is indeed the maximum found by Monte-Carlo simulation.

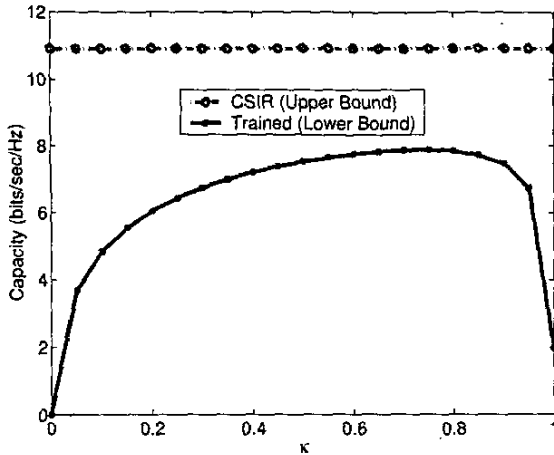


Fig. 1. Allocation of power between training and communication.

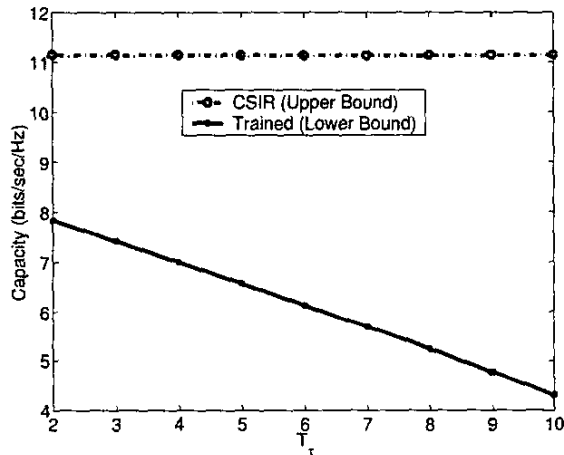


Fig. 2. The best training length.

Capacity as a function of the coding interval  $T$  is shown in the inset of Figure 3. The optimal power allocation,  $N = M = 2$  antennas, training interval of  $T_r = M$ , normalized Doppler frequency  $f = 0.003$  and specular parameter  $\beta = 0.5$  were used. The maximum is at  $T \approx 16$ . Monte Carlo evaluation of (31) in a similar manner was used to find the best  $\hat{T}$  for various values of  $\beta$ . These values are used in Figure 3 to plot the capacity as a function of  $\beta$  with the best training frequency  $\hat{T}$ , as well as with  $T = 13$ , which is the best training frequency for  $\beta = 1$ . Further research will investigate a reasonable choice for a training frequency to be used at all specular values. An initial choice will be the  $\hat{T}$  for a Rayleigh channel, since there does not seem to be much penalty for using it at all specular values, as illustrated in the figure.

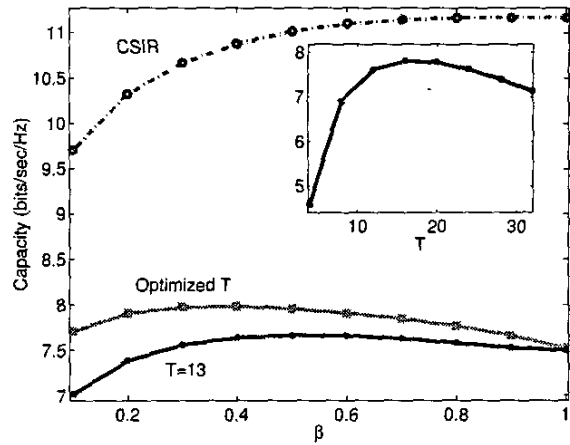


Fig. 3. Best training period.

## REFERENCES

- [1] B. Hassibi and B. M. Hochwald, "Optimal training in space-time systems," in *Proceedings Asilomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 743–747, 2000. Available from <http://mars.bell-labs.com/>.
- [2] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, pp. 139–157, May 1999. Available from <http://mars.bell-labs.com/>.
- [3] B. L. Hughes, "Differential space-time modulation," *IEEE Transactions on Information Theory*, vol. 46, pp. 2567–2578, November 2000. See <http://www4.ncsu.edu/blhughes/>.
- [4] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Transactions on Information Theory*, vol. 47, pp. 2335–2367, September 2001. Available from <http://mars.bell-labs.com/>.
- [5] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multi-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, pp. 359–383, February 2002. See <http://degas.eecs.berkeley.edu/dtse/pub.html>.
- [6] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath, "Optimal training for frequency-selective fading channels," in *Proceedings of the 2001 Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2105–2108, 2001.
- [7] C. B. Peel and A. L. Swindlehurst, "Performance of unitary space-time modulation in a continuously changing channel," in *Proceedings of the 2001 International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2433–2436, 2001. Available from <http://www.ee.byu.edu/peel>.
- [8] M. Godavarti, T. L. Marzetta, and S. S. (Shitz), "Capacity of a mobile multiple-antenna wireless link with isotropically random Rician fading," in *Proceedings of the 2001 IEEE International Symposium on Information Theory*, June 2001.
- [9] C. B. Peel and A. L. Swindlehurst, "Performance of space-time modulation for a general time-varying Rician channel model," *IEEE Transactions on Wireless Communications*, Submitted, 2002.
- [10] W. C. Jakes, *Microwave Mobile Communications*. IEEE Press, 1993.
- [11] C. B. Peel and A. L. Swindlehurst, "Capacity-optimal training for space-time modulation in time varying fading," in *The 2003 International Conference on Communications*, submitted, 2003.