PROBABILITY OF ERROR FOR TRAINED UNITARY SPACE-TIME MODULATION OVER A GAUSS-INNOVATIONS RICIAN 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 pairwise probability of error for trained unitary space-time modulation over channels with a constant specular component and time-varying diffuse fading is derived in this paper. We consider the case where the diffuse channel varies from sample to sample within a symbol according to a first-order AR model. Our previous results are reviewed which show that the effect of the time-varying diffuse channel can be described by an effective SNR that decreases with time. We derive pairwise probability of error expressions using these effective SNR values, which are shown by simulation to accurately describe performance.**

I. INTRODUCTION

The use of multiple transmit and receive antennas in wireless systems has received significant attention lately because of the high data rates they potentially offer [1], [2]. Though initial work has dealt with the situation where the receiver (and possibly the transmitter) knows the channel between each transmit and receive antenna, more recent work has focused on the case where to begin with neither the receiver nor the transmitter possess channel state information [3], [4], [5]. This latter scenario is usually dealt with using trained modulation, assuming that the channel is constant between training periods.

All of the above papers assume a piecewise-constant model of the time-varying channel coefficients. Such a model accurately describes the way a channel might appear in a time-division multiple access or frequency-hopping system, and its effects are simple to analyze. However, its inability to account for the memory of the channel make it less attractive in other applications. In addition, the fact that the time-varying channel is not truly piece-wise constant leads to significant modeling errors. In this paper, we examine via a probability of error expression the performance penalty incurred for multiple-antenna trained modulation when the piece-wise constant assumption is violated as the channel changes from sample to sample.

To incorporate channel time variation into our analysis, we use multiple first-order Gauss-innovations models to describe the time evolution of the channel coefficients. The model provides a reasonable fit to the temporal properties of physical channel models (such as Jakes' [6]) as will be illustrated with simulation

results. Our analysis focuses on the performance of trained space-time modulation when using unitary code matrices. After reviewing results from previous work [7], [8] which indicates that the effect of the Gauss-innovations model can be described by an effective SNR that decreases with time, we obtain pairwise probability of error expressions for this time-varying Rician channel model.

II. CHANNEL MODEL

Throughout the paper, we let $\mathcal{CN}(0,1)$ denote a zero-mean, unit-variance, circularly symmetric complex Gaussian distribution. The Frobenius norm will be denoted by $\|\cdot\|_F$, the expectation operator by $E[\cdot]$, and the determinant by $|\cdot|$.

A. Fading Channel Model

Our focus is on a flat-fading communications environment with M transmit and N receive antennas, with signal transmission occurring over T time instants. A complex channel coefficient 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, but not temporally white. At each receive antenna, interference and other disturbances add temporally and spatially independent noise to the signal.

We formalize these statements as follows: for $m = 1, \ldots, M$ transmit, and $n = 1, \ldots, N$ receive antennas, at time instants $t = 0, 1, \ldots, T-1$, the channel coefficient is $h_{m,n,t}$, with the signal transmitted from antenna m at time t denoted by $s_{t,m}$. We assume that the $M \times N$ matrix H_t formed from $h_{m,n,t}$ is normalized so that $E[\|H_t\|_F^2] = MN$, and the $T \times M$ matrix S formed from $s_{t,m}$ is normalized so that $E[\|S\|_F^2] = TM$. With these definitions, the data at receive antenna n is written

$$x_{t,n} = \sqrt{\frac{\rho}{M}} \sum_{m=1}^{M} h_{m,n,t} s_{t,m} + w_{t,n}, \qquad (1)$$

where we assume that the noise $w_{t,n}$ is $\mathcal{CN}(0,1)$. Due to the normalizations defined above, ρ represents the SNR expected at each receive antenna and does not depend on the number of transmit antennas.

In the case where the channel is constant ($H_t = H$, for $t = 0, 1, \ldots, T-1$) then (1) reduces to the piecewise-constant model of [4]:

$$X = \sqrt{\frac{\rho}{M}} SH + W. \qquad (2)$$

Though this quasi-static channel model is theoretically attractive, it is not always realistic, especially for environments with rapidly moving users. In such situations (1) is more applicable. It is shown in [8] that (1) can be written in a form similar to (2) with the addition of a diagonal matrix modifying the signal strength at each time instant. This analysis is done using a Gauss-innovations model (see Section II-C) to describe the time evolution of the channel.

B. Specular and Diffuse Channel Components

In our analysis we will separate the specular and diffuse components of the channel as follows:

$$H_t = \sqrt{1-\beta}H^s + \sqrt{\beta}H_t^d , \qquad (3)$$

where the specular part H^s is assumed to be known and time-invariant, and the elements of the diffuse component are modeled as $h_{m,n,t}^d \sim \mathcal{CN}(0,1)$. The only restriction on H^s is a power constraint $E[\|H^s\|_F^2] = MN$, which maintains the relationship in (1). We also decompose the signal power as $\rho = \rho^s + \rho^d$, where

$$\begin{aligned} \rho^s &= (1-\beta)\rho & (4) \\ \rho^d &= \beta\rho , & (5) \end{aligned}$$

and the parameter β allows tuning between a fully specular channel ($\beta = 0$) and a Rayleigh channel ($\beta = 1$). In practice we expect the channel to be composed of both diffuse and specular components, in which case $0 < \beta < 1$.

The probability of error expressions below assume a rank-one specular component; in this case H^s is written as the outer product of two fixed, but isotropically distributed random unit vectors [3]:

$$H^s = \sqrt{MN}\mathbf{uv}^H . \qquad (6)$$

C. A Gauss-Innovations Fading Channel Model

In Section III we characterize the performance of space-time modulation with the assumption that the current channel H_{r+t} occurs t samples after a reference (or estimated) channel H_r. We assume that between time r and $t+r$ the dispersive component of the channel H_{r+t}^d varies according to the following first-order auto-regressive (AR) or *Gauss-innovations* model:

$$H_{r+t}^d = \sqrt{\alpha_t}H_r^d + \sqrt{1-\alpha_t}E_{r+t} , \qquad (7)$$

where H_r^d and E_{r+t} have i.i.d. $\mathcal{CN}(0,1)$ elements, E_{r+t} is independent from symbol to symbol and $0 \leq \alpha_t \leq 1$. Under this model, H_{t+r}^d also has zero-mean, unit variance Gaussian entries. It is important to note that the channel is not described by a single AR model, but rather with multiple first-order models, one for each time difference between the current sample and the reference channel.

A time-invariant channel is produced for $\alpha_t = 1$, while $\alpha_t = 0$ indicates a completely random time-varying channel. For trained modulation $t \approx KT$, and demodulation is based on a channel estimate obtained $K \geq 1$ symbols in the past. The parameter α_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 autocorrelation function of an element of H_t^d. Solving the Yule-Walker equations for α_t in the first-order AR process (7) we obtain

$$\alpha_t = \left[\frac{r_{hh}(t)}{r_{hh}(0)}\right]^2 , \qquad (8)$$

which provides a reasonable choice for α_t. For example, assuming Jakes' model of the land mobile fading channel [6], $r_{hh}(t) = J_0(2\pi ft)$, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, $f = f_dT_s$, f_d is the maximum Doppler frequency in

the fading environment, and T_s is the sampling period. Under this model (8) leads to

$$\alpha_t = J_0(2\pi tf)^2 . \qquad (9)$$

This Gauss-innovations model is an appropriate approximation when using the maximum-likelihood decoders of [4] that depend only on a single reference channel. This fact is borne out by the simulation results of Section IV, where excellent agreement is obtained with data generated according to Jakes' model, but analyzed with the Gauss-innovations model.

D. Channel Estimation

A training signal of length T_τ is sent at the beginning of a frame, after which data symbols of length T are sent. We will consider the maximum likelihood (ML) estimate of the channel:

$$\hat{H}_r = \sqrt{\frac{M}{\rho}}\left(S_r^H S_r\right)^{-1} S_r^H X_r , \qquad (10)$$

where S_r is the training signal, X_r is the received training data, and all parameters are assumed to be known except the diffuse component of the channel H_r^d. Assuming unitary signals, the ML estimate becomes:

$$\hat{H}_r^d = \sqrt{\frac{M}{\rho^d T_\tau}}S_r^H\left(X_r - \sqrt{\frac{\rho^s}{M}}S_r H^s\right) = H_r^d + \sqrt{\frac{M}{\rho^d T_\tau}}\hat{W}_r , \qquad (11)$$

where H_r^d and W_r are the diffuse part of the channel and the receiver noise, respectively, seen during training.

III. PERFORMANCE FOR TRAINED MODULATION

In this section we analyze the performance of space-time modulation for a time varying channel where the decoder assumes that it has perfect knowledge of the channel, although it uses an estimate based on training. We assume the ML channel estimate described above, and compare the resulting performance with that obtained using the exact channel. An analysis similar to that presented below applies to the channel tracking techniques presented in [9] or [10].

A. Effective SNR Matrices

We assume that an ML estimate of the channel is obtained at some reference time r, and that the channel is constant during training. Because of the temporal variation of the channel, the quality of this estimate degrades with time. The theorem below shows that this degradation can be described as an effective SNR that decreases with time until training occurs again.

Theorem 1: Given the channel model of Section II, the effect of the channel variation t time samples after the channel estimate is that of a time-varying effective SNR and is described by the following equation:

$$X_{r+t} = \sqrt{\frac{\rho^s}{M}}P_t^s SH_r^s + \sqrt{\frac{\rho^d}{M}}P_t^d SH_r^d + \tilde{W}_{r+t} , \qquad (12)$$

where

$$\begin{aligned} P_t^s &= \text{diag}\{b_0, b_1, \cdots, b_{T-1}\} \\ P_t^d &= \text{diag}\{p_0, p_1, \cdots, p_{T-1}\} \end{aligned}$$

are diagonal matrices with elements given by

$$b_m = \sqrt{\frac{1}{1 + \alpha_{t+m+R_p} + (1 - \alpha_{t+m+R_p})\rho^d}} , \qquad (13)$$

and

$$p_m = \sqrt{\frac{\alpha_{t+m+R_p}}{1 + \alpha_{t+m+R_p} + (1 - \alpha_{t+m+R_p})\rho^d}} , \qquad (14)$$

where $R_p = \frac{T+1}{2}$.

Proof: See [8]. ∎

As $\rho \to \infty$ we find that the ESNR at each time instant ceases to depend on ρ, and depends solely on the parameters of the time-varying channel:

$$\lim_{\rho \to \infty} \sqrt{\rho^s} P_t^s = \sqrt{\frac{1-\beta}{\beta}} [\Delta^2 (I_M - D_t)]^{-\frac{1}{2}} \qquad (15)$$

$$\lim_{\rho \to \infty} \sqrt{\rho^d} P_t^d = D_t^{\frac{1}{2}} [\Delta^2 (I_M - D_t)]^{-\frac{1}{2}} , \qquad (16)$$

where D_t is the diagonal matrix formed from $\alpha_{t+R_p}, \ldots, \alpha_{t+T-1+R_p}$. In this case it is the time-variation of the channel rather than the SNR that limits performance.

B. Probability of Error

Using the above results, we are able to derive probability of error expressions for the general channel model introduced previously. Our analysis applies to a wide range of space-time coding approaches, including the linear block coding schemes of [1], [11] as well as unitary modulation [4], [12], [13]. We focus now on modulation with unitary matrices, though similar results apply for the linear codes. We first derive the pairwise probability of error for the Rayleigh fading channel, and then consider the rank-one Rician case.

Theorem 2 (P_e for Trained modulation): Given the effective data model of (12) for a Rayleigh channel ($\beta = 1$), and assuming the ML decoder of [4] (which assumes the data model (2) and that the channel H is known):

$$\arg \max_{l \in \{0,1,\ldots,L-1\}} \left\| X - \sqrt{\frac{\rho}{M}} S_l H \right\|_F , \qquad (17)$$

the pairwise probability of error is

$$P_e = \frac{1}{2\pi j} \int_{-\infty - j\epsilon}^{\infty - j\epsilon} \frac{1}{\omega} \prod_{m=1}^{M} \left[1 + (\omega^2 + j\omega)\delta_m^2 \right.$$
$$\left. + (\omega^2 - j\omega)d_m^2 - 2\omega^2 d_m \delta_m \right]^{-N} d\omega , \qquad (18)$$

where δ_m are the singular values of $\sqrt{\rho \frac{T}{M}}(P_t S_1 - S_2)$, d_m are the values of the diagonal matrix $\sqrt{\rho \frac{T}{M}}(P_t - I)$, and $\epsilon = -\max\left\{\frac{\delta_m + d_m}{\delta_m - d_m}\right\}$.

Proof: We assume that there are two signal matrices in the constellation, and begin by analyzing the probability of error given that the signal S_1 is sent:

$$P_{e|1} = P\left(\left\| X - \sqrt{\rho \frac{T}{M}} S_2 H \right\|^2 < \left\| X - \sqrt{\rho \frac{T}{M}} S_1 H \right\|^2 \;\Big|\; S_1\right) .$$

Let $A = \sqrt{\rho \frac{T}{M}}(P_t S_1 - S_2)$ and $A = U\Delta V^H$ be its SVD. We can disregard U, since we may pre-multiply our received data by U^H without changing its distribution. Also, we may post-multiply our constellation by V without changing the probability of error [12]. Then, with $D \triangleq \sqrt{\rho \frac{T}{M}}(P_t - I)$, we have the following expression for P_e:

$$P(\operatorname{tr}\{H^H(\Delta^2 - D^2)H + H^H(\Delta - D)W + W^H(\Delta - D)H\} < 0) .$$

Because Δ and D do not depend on whether S_1 or S_2 was sent, we note that $P_{e|1} = P_{e|2} \triangleq P_e$. Now, let

$$Z = \begin{bmatrix} H^H & W^H \end{bmatrix} \begin{bmatrix} \Delta^2 - D^2 & \Delta - D \\ \Delta - D & \mathbf{0} \end{bmatrix} \begin{bmatrix} H \\ W \end{bmatrix} = Y^H R Y.$$

Using the identity

$$\begin{vmatrix} A & B \\ C & I \end{vmatrix} = |A - BC| \qquad (19)$$

and results from [8] with $R = I$, we write

$$\phi_r(\omega) = \left| I + j\omega(\Delta^2 - D^2) + \omega^2(\Delta - D)^2 \right|^{-N} \qquad (20)$$

$$= \prod_{i=1}^{M} \left[1 + j\omega(\delta_m^2 - d_m^2) + \omega^2(\delta_m - d_m)^2 \right]^{-N} \qquad (21)$$

We choose as our contour $C = (-\infty - j\epsilon, \infty - j\epsilon)$. The poles of $\phi_r(\omega)$ are all along the $j\omega$ axis; we want to choose $\epsilon > 0$ so that we may exchange the order of integration, but small enough that our contour of integration does not include any other poles. Choosing $\epsilon = -\max_m\left\{\frac{\delta_m + d_m}{\delta_m - d_m}\right\}$ satisfies these constraints. Using (21) we obtain the desired result. ∎

The integral in (18) can be easily evaluated using common numerical techniques.

It is somewhat more difficult to obtain results for a Rician channel; the methods described in [8] do not extend easily to a non-Gaussian distribution. Rather than an exact expression, we settle instead for an approximation assuming a rank-one specular component. We assume that the ESNR matrices P_t^s and P_t^d are known at the receiver, and can be used in the ML decoder; we also assume that diagonal signals are used. Under these assumptions the probability of error P_e is shown in [8] to be

$$\frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{1}{\omega^2 + 1/4} \prod_{m=1}^{M} \left[1 + (\omega^2 + \tfrac{1}{4})\rho\beta\frac{T}{M}(p_m^d d_m)^2 \right]^{-N} \cdot \qquad (22)$$
$$\exp\left\{ -N \sum_{m=1}^{M} \left(\frac{p_m^s}{p_m^d}\right)^2 \frac{1-\beta}{\beta} \frac{(\omega^2 + 1/4)\rho\beta T/M(p_m^d d_m)^2}{1 + (\omega^2 + 1/4)\rho\beta T/M(p_m^d d_m)^2} \right\} d\omega$$

where d_m are the singular values of $S_1 - S_2$. We will see in Section IV that (22) provides a good approximation.

IV. SIMULATION RESULTS

In Figure 1 we show simulation results using channel coefficients from a geometrical single-bounce model [14], which is similar to Jakes' model. We consider an uplink scenario where the base and mobile are separated by two thousand wavelengths (2000λ), a single base antenna is used, and two mobile transmitter antennas are separated by λ. The mobile is surrounded by a 200λ disk of twenty five randomly placed scatterers, and is moving with a normalized Doppler frequency of $f_d = 0.02$. Each data point in the figure is from one hundred experiments of 10^4 samples each. A diagonal signal constellation containing two unitary matrices was used. Results for simulations are shown with solid lines, while analytic results assuming the Gauss-innovations model are shown with dotted lines. Results from Theorem 2 for trained modulation are shown along with results for differential modulation from [8] for purposes of comparison. The excellent agreement between simulation and analysis lends support to the model and probability of error expressions presented in this paper.

Figure 2 presents probability of error performance as a function of the specular parameter β with a rank-one specular component. The simulation parameters were $M = 2$ transmit antennas, $N = 2$ receive antennas, a constellation of size $L = 2$, $\rho = 10$dB, and coefficients that obey Jakes' model with parameter $f_d = 0.003$.
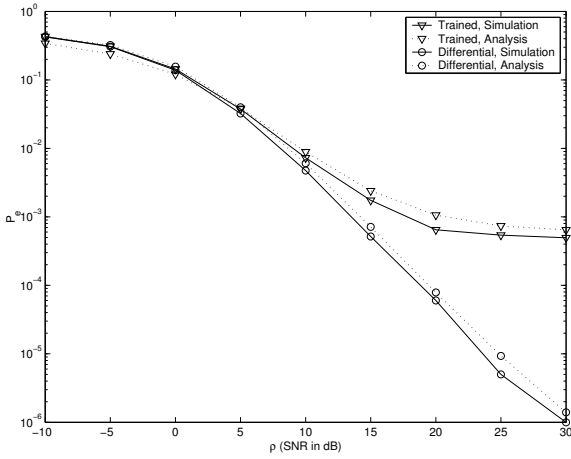
Fig. 1. Comparison of differential and trained modulation using channel data from a geometrical single-bounce model.
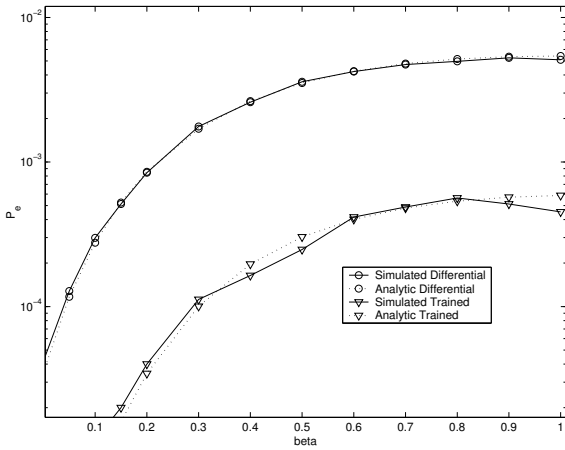

Fig. 2. Performance as a function of specular parameter β.

We show analytic and simulation results for trained modulation with perfect channel estimation and differential modulation; the analytic results are similar to those given in this paper and are found in [8]. In this scenario, the specular channel yields better performance than a diffuse channel.

Figure 3 shows the behavior of trained unitary modulation versus the length of the training interval. We let the training interval be $t = KM$, and vary K from 1 to 10. The fading parameter $f_d = 0.003$ was used in a fully diffuse ($\beta = 1$) channel with signal to noise ratio of $\rho = 20$dB. Simulations results are shown using results from Theorem 2 as well as for trained modulation with channel state information at the receiver (indicated with "CSIR" in the plots) and for differential modulation, using results from [8]. The analytic and simulation probability of error results agree well and show that probability of error increases linearly as K increases. This is not the entire picture, however, because the rate is increasing with K as well, according to $K/(K + 1)$. A detailed analysis of the training frequency K which maximizes capacity is found in [15].
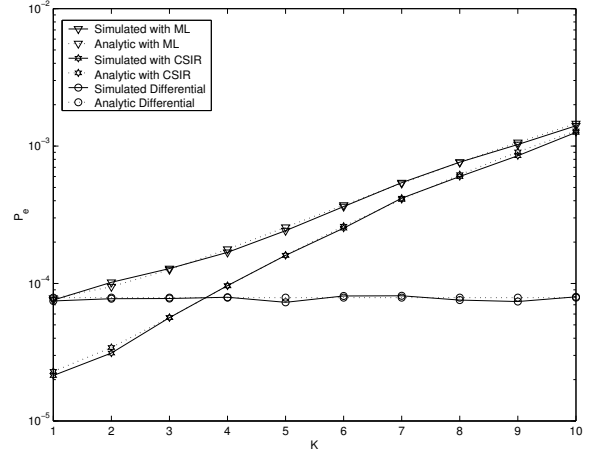

Fig. 3. Performance versus training period K for Differential and Trained Modulation.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Nov/Dec 1999.

[3] 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.

[4] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, pp. 543–564, March 2000.

[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.

[6] W. C. Jakes, *Microwave Mobile Communications*. IEEE Press, 1993.

[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.

[8] 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.

[9] Z. Liu, G. B. Giannakis, S. Zhou, and B. Muquet, "Space-time coding for broadband wireless communications," *Wireless Communications and Mobile Computing*, vol. 1, no. 1, pp. 35–36, 2001.

[10] R. Schober and L. H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Transactions on Communications*, vol. 50, pp. 768–777, May 2002.

[11] B. Hassibi and B. M. Hochwald, "High-rate linear codes," in *Proceedings of the 2001 International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2461–2464, 2001.

[12] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Transactions on Communications*, vol. 49, pp. 2041–2052, March 2000.

[13] B. L. Hughes, "Differential space-time modulation," *IEEE Transactions on Information Theory*, vol. 46, pp. 2567–2578, November 2000.

[14] T. Svantesson, *Antennas and Propogation from a Signal Processing Perspective*. PhD thesis, Chalmers University of Technology, 2001.

[15] 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.