Pairwise Probability of Error for Differential Space-Time Modulation over a Time-Varying 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 differential unitary space-time modulation for channels with a constant specular component and time-varying diffuse fading is derived in this paper. We consider the case where the channel varies from sample to sample within a symbol according to a first-order Gauss-innovations 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

Wireless communication systems must be able to effectively deal with unknown propagation channels that vary with time. A common way of handling this is differential modulation, which exploits the special group structure of the symbols to obviate the need for channel estimates. There are limitations to this technique; differential modulation assumes that the channel is constant for two successive symbols. In this paper, we examine via a probability of error expression the performance penalty incurred for multiple-antenna differential modulation when these assumptions are violated as the channel changes from one sample to the next.

Wireless systems with multiple transmit and receive antennas have 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 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

\[
x_{t,n} = \sqrt{\frac{\rho}{M}} \sum_{m=1}^{M} h_{m,n,t} s_{t,m} + w_{t,n} , \tag{1}
\]
where we assume that the noise $w_{t,n}$ is $CN(0,1)$. Due to the normalizations defined above $\rho$ 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]:

$$X = \sqrt{\frac{\rho}{M} SH + W}.$$

(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 [9] 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:

$$H_t = \sqrt{1 - \beta} H^s + \sqrt{\beta} H^d_t,$$

(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 $h^d_{m,n,t} \sim CN(0,1)$. The only restriction on $H^s$ is a power constraint $E[\|H^s\|^2] = MN$, which maintains the relationship in (1). We also decompose the signal power as $\rho = \rho^s + \rho^d$, where

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

(4)

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

Simulation results will be presented for a rank-one specular component; in this case $H^s$ is written as the outer product of two isotropically distributed random unit vectors [3]:

$$H^s = \sqrt{MN} u v^H.$$

(5)

A rank-one model similar to this has been used for analysis of capacity in [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 $H^d_{r+t}$ varies according to the following first-order auto-regressive (AR) or Gauss-innovations model:

$$H^d_{r+t} = \sqrt{\alpha_t} H^d_r + \sqrt{1 - \alpha_t} E_{r+t},$$

(6)

where $H^d_r$ and $E_{r+t}$ have i.i.d. $CN(0,1)$ elements, $E_{r+t}$ is independent from symbol to symbol and $0 \leq \alpha_t \leq 1$. Under this model, $H^d_{r+t}$ 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. With differential coding, $t = M$ and demodulation is based on the previous symbol (of length $M$). On the other hand, for trained modulation $t \approx KT$, and demodulation is based on a channel estimate obtained $K \geq 1$ symbols in the past.

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 autocorrelation function of an element of $H^d_t$. Solving the Yule-Walker equations for $\alpha_t$ in the first-order AR process (6) we obtain

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

(7)

which provides a reasonable choice for $\alpha_t$. For example, assuming Jakes’ model of the land mobile fading channel [7], $r_{hh}(t) = J_0(2\pi ft)$, 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 model (7) leads to

$$\alpha_t = J_0(2\pi ft)^2.$$

(8)

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.

III. PROBABILITY OF ERROR FOR DIFFERENTIAL MODULATION

A. Differential Modulation

Differential space-time modulation [10], [11] applies to the case of an unknown channel that is constant over each pair of consecutive transmitted symbols. Differential encoding rotates the previous unitary space-time transmitted data $S_{r-1}$ by the current space-time symbol $V_r$ to obtain the current data to transmit: $S_r = V_r S_{r-1}$, where $\tau$ is the symbol index. Using these definitions, the following expression for the current received data $X_r$ is obtained when the channel is time-invariant:

$$X_r = V_r X_{r-1} + \sqrt{2\tilde{W}_r},$$

(9)

where the elements of $\tilde{W}_r$ are i.i.d. $CN(0,1)$ variables. Because the effective channel $(X_{r-1})$ has signal strength $\rho$, the system has an effective SNR of $\rho/2$. This factor of two corresponds to the well-known 3 dB loss in performance incurred when using DPSK rather than coherent PSK. With the identification of $X_{r-1}$ as the effective channel, (9) is simply (2) with half the signal strength that would be seen with coherent detection.

2699
As noted in [9], trained modulation has a performance ceiling at high SNR which decreases with the length of the symbol and the training period. It is useful to compare differential modulation with trained modulation implemented with the shortest possible training period. We thus consider the case where known pilot symbols are transmitted every other symbol. Using these unitary space-time symbols, we note that the ML estimate for the channel assuming (2) is
\[
\hat{H} = \sqrt{\frac{M}{\rho}} S_T^H X_T = H + \sqrt{\frac{M}{\rho}} \tilde{W}_T,
\]
where \( S_T \) is a (unitary) matrix of training data and \( \tilde{W}_T \) has i.i.d. \( \mathcal{CN}(0,1) \) elements. With these definitions, we obtain an equation for the effective channel seen by the received data:
\[
X = \sqrt{\frac{\rho}{M}} S H + \tilde{W}_T + W = \sqrt{\frac{\rho}{M}} S H + \sqrt{2} \tilde{W}
\]
where \( \tilde{W} \) has the same distribution as \( W \). The same 3 dB penalty in SNR is obtained as with differential modulation; in addition, we will show below that approximately the same error floor is obtained at high SNR. The obvious problem with training this frequently is that the data transmission rate is half that of differential modulation. A more reasonable approach would be, for example, to send training data every tenth symbol. However, as we show below, even though the rate would be 90% of that for differential modulation, a much worse error floor at high SNR will result.

Recall that for differential modulation, the previous received data block serves as the effective channel, and thus it makes sense to use the channel seen by the previous symbol as a reference. This does not give a unique solution, however, because there is a different channel seen by the previous symbol at each of the \( T \) time instants. We desire to place the reference channel in the position which is “closest” to the other channels seen by the previous symbol. We choose to use as reference the channel that is temporally in the center of the previous symbol; the first time sample in the current symbol will thus be \( \frac{T-1}{2} + 1 \) time samples removed from the reference.

**Theorem 1:** Assuming that the channel obeys the model in Section II, the effect of the time-varying diffuse channel on differential modulation is that of a time-varying and decreasing effective SNR. It is described by the following equation:
\[
X_{r+t} = \sqrt{\frac{\rho}{M}} P^s_t S H^s_r + \sqrt{\frac{\rho d}{M}} \tilde{P} S H^d_r + \tilde{W}_{r+t},
\]
where \( P^s_t \) = \( \text{diag}(b_0, b_1, \ldots, b_{T-1}) \) and \( \tilde{P} \) = \( \text{diag}(p_0, p_1, \ldots, p_{T-1}) \) are diagonal matrices with elements given by
\[
b_m = \sqrt{\frac{1}{1 + \alpha_{m+R_P} + (1 - \alpha_{m+R_P}) p_m^2}} \leq \frac{1}{\sqrt{2}}, \quad (12)
\]
\[
p_m = \sqrt{\alpha_{m+R_P} b_m} \leq \frac{1}{\sqrt{2}}. \quad (13)
\]

**Proof:** See [9].

### B. Characteristic Function for the Trace of Quadratic Forms

Before deriving the probability of error, we will review the characteristic function of the trace of a matrix quadratic form. This is an extension of results in [12] from quadratic forms involving Gaussian-distributed vectors to those involving matrices. This technique encompasses those used in [4], [13], [14].

Let \( Z = Y^H Q Y \), where \( Y \) is an \( M \times N \) matrix, and where the \( n \)th column \( y_n \) is distributed as \( \mathcal{CN}(\mathbf{y}_n, \mathbf{R}) \). The characteristic function for \( r = \text{tr}(Z) \) is
\[
\phi_r(\omega) = |I + j\omega R|^{-N} \exp \left( -\text{tr} \{ Y^H R^{-1} [I - (I + j\omega R)^{-1}] Y \} \right), \quad (14)
\]
where \( \tilde{Y} = E[Y] \). This can be shown using the characteristic function of the \( N \) terms \( z_{n} \) as follows:
\[
\phi_r(\omega) = \prod_{n=1}^{N} \phi_{z_{n}}(\omega)
\]
which is half that of differential modulation. A more reasonable approach would be, for example, to send training data every tenth symbol. However, as we show below, even though the rate would be 90% of that for differential modulation, a much worse error floor at high SNR will result.

Recall that for differential modulation, the previous received data block serves as the effective channel, and thus it makes sense to use the channel seen by the previous symbol as a reference. This does not give a unique solution, however, because there is a different channel seen by the previous symbol at each of the \( T \) time instants. We desire to place the reference channel in the position which is “closest” to the other channels seen by the previous symbol. We choose to use as reference the channel that is temporally in the center of the previous symbol; the first time sample in the current symbol will thus be \( \frac{T-1}{2} + 1 \) time samples removed from the reference.

**Theorem 1:** Assuming that the channel obeys the model in Section II, the effect of the time-varying diffuse channel on differential modulation is that of a time-varying and decreasing effective SNR. It is described by the following equation:
\[
X_{r+t} = \sqrt{\frac{\rho}{M}} P^s_t S H^s_r + \sqrt{\frac{\rho d}{M}} \tilde{P} S H^d_r + \tilde{W}_{r+t},
\]
where \( P^s_t \) = \( \text{diag}(b_0, b_1, \ldots, b_{T-1}) \) and \( \tilde{P} \) = \( \text{diag}(p_0, p_1, \ldots, p_{T-1}) \) are diagonal matrices with elements given by
\[
b_m = \sqrt{\frac{1}{1 + \alpha_{m+R_P} + (1 - \alpha_{m+R_P}) p_m^2}} \leq \frac{1}{\sqrt{2}}, \quad (12)
\]
\[
p_m = \sqrt{\alpha_{m+R_P} b_m} \leq \frac{1}{\sqrt{2}}. \quad (13)
\]

**Proof:** See [9].
where

\[
R = \begin{bmatrix}
S_1^H R X S_1 & S_1^H R X S_2 \\
S_2^H R X S_1 & S_2^H R X S_2
\end{bmatrix}
\] (23a)

\[
Q = \begin{bmatrix}
-I & 0 \\
0 & I
\end{bmatrix}
\] (23b)

\[
R_X = 2\rho P^d S S^H P^d + I
\] (23c)

\[
\bar{Y} = \sqrt{2\rho} \begin{bmatrix}
S_1^H P^s S_1 H^s \\
S_2^H P^s S_1 H^s
\end{bmatrix}
\] (23d)

**Proof:** Using Theorem 1, we write

\[
X = \bar{X} + \sqrt{\frac{\rho^d}{M}} P^d S H^d + W,
\]

where \(\bar{X} = \sqrt{\rho^d} P^s S H^s\) is due to the time-invariant specular component. Each column of \(X\) has covariance

\[
R_X = \frac{\rho^d}{M} P^d S S^H P^d + I.
\]

The probability that \(S_2\) is decoded when \(S_1\) is sent is

\[
P_{e|1} = P(\|S_2 X\|^2 > \|S_1 X\|^2 | S_1) = P(\text{tr} \{X^H (S_2 S_2^H - S_1 S_1^H) X\} > 0 | S_1) = P(\text{tr}(Y^H Q Y) > 0 | S_1),
\]

where

\[
Q = \begin{bmatrix}
-I & 0 \\
0 & I
\end{bmatrix}
\]

\[
Y = \begin{bmatrix}
Y_1 \\
Y_2
\end{bmatrix} = \begin{bmatrix}
S_1^H X \\
S_2^H X
\end{bmatrix}.
\]

Each column of \(Y\) has covariance

\[
R = E [(Y_n - \bar{Y}_n)(Y - \bar{Y})^H] = \begin{bmatrix}
S_1^H R X S_1 & S_1^H R X S_2 \\
S_2^H R X S_1 & S_2^H R X S_2
\end{bmatrix},
\]

where the mean of \(Y\) is

\[
\bar{Y} = \sqrt{\frac{\rho^d}{M}} \begin{bmatrix}
S_1^H P^s S_1 H^s \\
S_2^H P^s S_1 H^s
\end{bmatrix}.
\]

We may now use the results of (14) to find the characteristic function of \(r = \text{tr}(Y^H Q Y)\) and the probability of error, which differs by a sign from that in (19):

\[
P(r > 0) = \frac{-1}{2\pi j} \int_{-\infty - j\epsilon}^{\infty - j\epsilon} \frac{1}{\phi_r(\omega)} d\omega.
\]

Using the fact that for differential modulation the coding effectively takes place over two symbols at a time (\(\frac{T}{M} = 2\)), we have the results in (22).

A similar technique could be used to generate probability of error expressions for the signals designed for the quasi-static, unknown channel, including those in [4], [15], [16], while the expression above applies only for differential modulation [11], [10].

Theorem 2, in fact, requires the differential signals to be in a slightly different form than presented earlier. For a signal \(V\) from a constellation of unitary signals, the signal to be used in the theorem is

\[
S = \frac{1}{\sqrt{2}} \begin{bmatrix}
I_M \\
0
\end{bmatrix}.
\]

In addition, given a square ESNR matrix \(P\) from Theorem 1 the matrix to be used in Theorem 2 is

\[
\bar{P} = \begin{bmatrix}
P & 0_{M \times M} \\
0_{M \times M} & P
\end{bmatrix}.
\]

**IV. SIMULATION RESULTS**

We begin by presenting results which compare analytic and simulation results for differential modulation with one, two and three transmit antennas as a function of SNR. Figure 1 shows performance with the diagonal codes from [11] with \(N = 1\) receive antenna, a constellation of size \(L = 2\), fading parameter \(f = 0.025\), and specular parameter \(\beta = 0.5\). One hundred trials of ten thousand samples each were averaged to produce the results in this and the following simulations. The analytic and simulation results agree very well. Note that each curve has an error floor at high SNR due to the time-variation of the channel.

In Figure 2 we show simulation results using channel coefficients from a geometrical single-bounce model [17], which is similar to Jakes’ model. We consider an uplink scenario where the base and mobile are separated by two thousand wavelengths, a single base antenna is used, and two mobile transmitter antennas are separated by a wavelength. The mobile is surrounded by a disk of twenty five randomly placed scatterers, with radius of two hundred wavelengths and the mobile moving according to 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 with are shown with the solid lines, while analytic results assuming the Gauss-innovations model are shown with dashed lines. Results from [9] for
trained modulation with ML channel estimation are shown for comparison. The excellent agreement between simulation and analysis lends support to the model and probability of error expressions presented in this paper.

Figure 3 presents probability of error performance as a function of the specular parameter $\beta$ when used 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 $f_d = 0.003$. We show analytic and simulation results for differential modulation and trained modulation with genie-aided channel estimation. In this scenario, the specular channel gives better performance than a diffuse channel.

REFERENCES