PERFORMANCE ANALYSIS OF BLIND ADAPTIVE MIMO RECEIVERS

B. Balasingam†, M. Bolic†, S. Shahbazpanahi‡ and T. Kirubarajan*

†School of Information Technology and Engineering, University of Ottawa, Canada
‡Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Canada
Department of Electrical and Computer Engineering, McMaster University, Canada
E-mails: †{balasing,mbolic}@site.uottawa.ca
‡shahram.shahbazpanahi@uoit.ca, *kiruba@mcmaster.ca

1. INTRODUCTION

Multiple-input multiple-output (MIMO) systems combined with orthogonal space time block codes (OSTBCs) and orthogonal frequency division multiplexing (OFDM), known as MIMO-OFDM in the literature, are playing an important role in current and future wireless communications.

In this paper, we derive approximate BER expressions for the blind adaptive MIMO-OFDM receivers. Assuming the knowledge of the initial channel up to a priori probability distribution, we have derived the instantaneous signal to interference and noise ratio (SINR) for consecutive transmission blocks in the absence of training, by exploiting Kalman filtering to track the channel. Unlike previous methods such as [1] the derivation is general for any number of transmitting and receiving antennas, for general definition of OSTBCs and for any number of sub-carriers. The derived analysis can be used to study the effect in performance for different types of the channel estimation techniques (or stages) such as the simple Kalman predicted channel estimate (also known as 1 step prediction method), Kalman updated channel [2] and the steady-state channel estimate [2] etc. Further, using the resulting expression for BER, we are able to study the effect in performance due to the number of transmitting/receiving antennas, number of sub-carriers, mobile velocity and the amount of assumed channel length.

2. BACKGROUND

2.1. Orthogonal space-time coded OFDM systems

Consider a MIMO-OFDM system with L sub-carriers, N transmit antennas and M receive antennas. Using vector notations, the received signal vector for the rth tone during the nth OFDM word can be written as

\[ y_r[n] = X_r[n]H_i[n] + v_r[n] \]  

(1)

where \( y_r[n] \) is the \( M \times 1 \) vector of received signals, \( x_r[n] \) is the \( N \times 1 \) vector of transmitted signals, and \( v_r[n] \) is the \( M \times 1 \) vector of noise at the rth tone. The noise is assumed to be uncorrelated zero-mean complex Gaussian with variance \( \sigma^2_v/2 \) per real dimension. The \( N \times M \) complex matrix \( H_i[n] \) is the frequency response of the MIMO channel at the rth sub-carrier and can be written, in terms of the time-domain channel impulse response, as

\[ H_i[n] = \sum_{p=0}^{P-1} G_p[n]e^{-j2\pi ip/L} \]  

(2)

where \( j = \sqrt{-1} \), the \((r,q)\)th element of \( G_p[n] \) is the \( p \)th tap of the time-domain channel impulse response from the \( r \)th transmitter to the \( q \)th receiver, and \( P \) is the maximum delay spread of the channel. We assume that \( P < L \). We consider a block transmission scheme with block length \( T \) and assume that within the OFDM word length \( PT \), the channel frequency response is fixed. However, between different OFDM words the channel frequency response can change. Based on such an assumption, the \( n \)th received block for the \( r \)th sub-carrier can be written as

\[ Y_r[n] = X_r[n]H_i[n] + V_r[n] \]  

(3)

where

\[ Y_r[n] = [y_r^T[nT - T + 1], \ldots, y_r^T[nT]]^T \]

\[ X_r[n] = [x_r^T[nT - T + 1], \ldots, x_r^T[nT]]^T \]

\[ V_r[n] = [v_r^T[nT - T + 1], \ldots, v_r^T[nT]]^T \]

\( Y_r[n] \) denote the received data, the transmitted data, and the measurement noise matrices, respectively, and \((\cdot)^T\) represents the transpose operator.

In space-time block coding, the matrix \( X_r[n] \) is a mapping that transforms a block of complex symbols into a \( T \times N \) complex matrix. Hence, we hereafter replace \( X_r[n] \) with \( X(s_i[n]) \) where \( s_i[n] \) of length \( K \) is the \( i \)th transmitted symbol vector of the \( r \)th OFDM word. Let us define \( s_i[n] \) as

\[ s_i[n] = (s_{1,i}[n], s_{2,i}[n], \ldots, s_{K,i}[n])^T \]

By selecting \( X(s_i[n]) \) as an OSTBC [3,4], we can write the observed data of the \( i \)th sub-carrier as [4]

\[ \hat{y}_{i,n} \triangleq \hat{Y}_i[n] = A(h_i[n])\hat{s}_{i,n} + \hat{v}_{i,n} \]  

(4)
where \( \mathbf{h}_i[n] = \mathbf{H}_i[n] \), \( \mathbf{s}_{n,i} \), \( \tilde{\mathbf{s}}_{n,i} \), \( \tilde{\mathbf{v}}_{n,i} = \mathbf{V}_i[n] \), the “underline” operator for a matrix \( \mathbf{P} \) is defined as \( \underrightarrow{\mathbf{P}} \triangleq \{\Re(\mathbf{P})\}^T \{\Im(\mathbf{P})\}^T \) and \( \mathrm{vec} \{\cdot\} \) refers to the vectorization operator stacking all the columns of a matrix on top of each other, and \( \Re(\cdot) \) and \( \Im(\cdot) \) represent the real and imaginary parts, respectively. The matrix \( \mathbf{A}(\cdot) \) is the so-called virtual channel matrix and for any real \( 2MN \times 1 \) vector \( \mathbf{h}_i[n] = \mathbf{H}_i[n] \) is given by [5]

\[
\mathbf{A}(\mathbf{h}_i[n]) = \left[ \mathbf{X}(\mathbf{e}_1) \mathbf{H}_n[,1] \cdots \mathbf{X}(\mathbf{e}_K) \mathbf{H}_n[,K] \right] = \left[ \mathbf{X}(\mathbf{e}_1) \mathbf{H}_n[,1] \mathbf{H}_n[,2] \cdots \mathbf{X}(\mathbf{e}_K) \mathbf{H}_n[,K] \right]
\]

where \( \mathbf{e}_k \) is the \( k \)-th column of the \( K \times K \) identity matrix \( \mathbf{I}_K \).

### 2.2. Conditional ML decoding

Using (4), the \( n \)-th received OFDM word can be written as

\[
\mathbf{y}_n = \mathbf{A}(\mathbf{h}_i[n]) \mathbf{s}_{n,i} + \mathbf{v}_n
\]

where \( \mathbf{y}_n = [\mathbf{y}_{n,0}, \ldots, \mathbf{y}_{n,L-1}]^T \), \( \mathbf{s}_{n,i} = [\mathbf{s}_{n,0}, \ldots, \mathbf{s}_{n,L-1}]^T \), \( \mathbf{A}(\mathbf{h}_i[n]) = \{\mathbf{A}(\mathbf{h}_i[n],[1]), \ldots, \mathbf{A}(\mathbf{h}_i[n],[K])\} \), \( \mathbf{h}_i[n] = [h_{0,n}, \ldots, h_{2MN,n}]^T \), and \( \mathbf{v}_n = [\mathbf{v}_{n,0}, \ldots, \mathbf{v}_{n,L-1}]^T \).

Given the channel vector \( \mathbf{h}_i[n] \), the optimal maximum likelihood (ML) detection for OSTBC based MIMO-OFDM system can be done in a tone-by-tone basis. The linear receiver computes \( \mathbf{\hat{s}}_{n,i} \), the estimate of \( \mathbf{s}_{n,i} \), as

\[
\mathbf{\hat{s}}_{n,i} = \frac{1}{\|\mathbf{h}_i[n]\|^2} \mathbf{A}^T(\mathbf{h}_i[n]) \mathbf{y}_{n,i}.
\]

The symbol-by-symbol decoder then builds \( \mathbf{\tilde{s}}_{n,i} \), the estimate of vector \( \mathbf{s}_{n,i} \), as \( \mathbf{\tilde{s}}_{n,i} = [\mathbf{I}_K \otimes \mathbf{\hat{\mathbf{s}}}_{n,i}] \). The \( k \)-th element of \( \mathbf{\tilde{s}}_{n,i} \) is compared with all the points in the constellation corresponding to \( \mathbf{s}_{n,i} \) and the closest point to the \( k \)-th element of \( \mathbf{\tilde{s}}_{n,i} \) is accepted as the \( k \)-th decoded symbol.

### 2.3. Channel modeling

For each tap \( p \), we model the channel variation between adjacent OFDM words as a first order autoregressive (AR) model. It was shown in [6] that the first order AR model provides a sufficiently approximate channel model for time-selective channels. Hence,

\[
\mathbf{G}_p[n+1] = \alpha_p \mathbf{G}_p[n] + \mathbf{U}_p[n] \quad p = 0, \ldots, P - 1
\]

where elements of \( \mathbf{U}_p[n] \) are i.i.d Gaussian with zero mean and variance \( \sigma_p^2 \) and \( \sigma_p^2 = \sigma_p^2(1 - \alpha_p^2) \). The complex parameters \( \alpha_p \), \( p = 0, 1, \ldots, P - 1 \) can be estimated [7] and hence are assumed known here.

As noted in (2), the frequency response of the channel for the \( i \)-th sub-carrier, \( \mathbf{h}_i[n] \), is a function of the time domain MIMO channel impulse response \( \{\mathbf{G}_p[n]\}_{p=0}^{P-1} \). Usually, the channel taps are much smaller in length compared to the number of tones. Hence, it is desirable to track the channel in time domain for better performance.

We now use the relationship between the channel frequency response and the channel impulse response to represent the observed data (6) in terms of the channel impulse response. Later, this model will be used to derive the Kalman filtering based tracking scheme for the channel coefficients.

Using the Fourier relationship between the channel frequency response \( \mathbf{h}_i[n] \) and channel impulse response \( \mathbf{g}[n] \) we can write \( \mathbf{h}_i[n] = \mathbf{W}_i \mathbf{g}[n] \) where

\[
\mathbf{W}_i = \begin{bmatrix}
\Re\{\mathbf{w}_i^T\} \otimes \mathbf{I}_{MN} & -\Im\{\mathbf{w}_i^T\} \otimes \mathbf{I}_{MN} \\
\Im\{\mathbf{w}_i^T\} \otimes \mathbf{I}_{MN} & \Re\{\mathbf{w}_i^T\} \otimes \mathbf{I}_{MN}
\end{bmatrix}.
\]

Here, \( \otimes \) is the Kronecker product, \( \mathbf{g}[n] = \mathbf{G}[n], \mathbf{G}[n] = (\mathbf{g}[n] \ldots \mathbf{g}[P-1,n]) \), \( \mathbf{w}_i = \{1, e^{j2\pi i/L}, \ldots, e^{j2\pi (P-1)/L}\} \).

### 3. PERFORMANCE ANALYSIS

Let us assume that we have the knowledge of the channel at the \((n-1)\)-th block in the form of mean and covariance as follows:

\[
\mathbf{g}[n-1|n-1] = \hat{\mathbf{g}}_{n-1}
\]

\[
\mathbf{P}[n-1|n-1] = \delta_{n-1} \mathbf{I}_{2MN}
\]
In the absence of such knowledge pilot symbols could be employed at the \((n - 1)\)th block to initialize the channel. For simplicity let us assume that the channel statistics remain the same at each taps, i.e.,
\[ a_0 = a_1 = \ldots = a_p = a. \]

The initial estimate and the corresponding covariance in (13)-(14) can be transformed to frequency domain for the \(s\)th sub-carrier by pre-multiplying by \(W_s\):

\[
\begin{align*}
\hat{h}_s[n-1] & = W_s g[n-1] \\
P_s[n-1] & = W_s P[n-1 | n-1 | W_s^T = \delta_{s-1}^n I_{2MNP} \\
\end{align*}
\]

(15)

(16)

where \(\delta_{s-1}^n = P_{\beta_{s-1}^n} \).

Without any pilot available at block \(n\), the Kalman prediction of the channel at block \(n\) is written as

\[
\begin{align*}
g[n-1] & = F g[n-1 | n-1] \\
& = g[n] + \varepsilon [n | n-1] \\
\end{align*}
\]

where \(\varepsilon [n]\) is the true channel and \(\varepsilon [n-1 | n-1] = 2MN \times 1\) vector that accounts for the channel estimation errors at block \(n\). Further, it can be verified that \(\varepsilon [n | n-1]\) has zero mean and covariance matrix

\[
P_{\varepsilon}[n-1 | n-1] = \mathbb{E}P [n-1 | n-1 | F^T] + Q = \beta_{s}^n I_{2MNP}
\]

(17)

(18)

where \(\beta_{s}^n = \delta_{s-1}^n \alpha^2 + \sigma_{\eta}^2 / 2\).

The channel predictions described in (17)-(18) is transformed to the frequency domain for the \(s\)th sub-carrier by pre-multiplying by \(W_s\):

\[
\begin{align*}
h_s[n-1] & = h_s[n] + \xi [n | n-1] \\
P_s[n-1] & = W_s P_{s}[n-1 | n-1] W_s^T = \beta_{s}^n I_{2MNP} \\
\end{align*}
\]

(19)

(20)

where \(h_s[n]\) is the true response, \(\xi [n | n-1]\) is the error corresponding to \(\varepsilon [n | n-1]\) and \(\beta_{s}^n = P_{\beta_{s-1}^n} \). It can be noticed that \(\xi [n | n-1]\) has zero mean and covariance matrix \(P_{\xi}[n-1 | n-1]\). It is worth mentioning that \(h_s[n-1 | n-1]\) is zero mean with covariance \(P_{h}[n-1 | n-1] = \mathbb{E} [h_s[n-1 | n-1] h_s[n-1 | n-1]^T] = \rho_{\eta} I_{2MNP}\) where \(\rho_{\eta} = \rho_{\alpha} (\frac{\sigma_{\eta}^2}{2\sigma_{\alpha}^2 + \sigma_{\eta}^2})\).

Using the predicted channel \(h_s[n-1]\) the data could be decoded either by employing a decision directed technique or through the blind detection technique reported in [9]. Assuming that the data was decoded error free, the Kalman updated channel is given as

\[
g[n | n] = g[n | n-1] + G_K \nu[n]
\]

(21)

where \(G_K\) is the innovation process that is zero-mean with covariance matrix \(P_s[n]\) (see 2.4). Similar to (17) and 18, we can write the above Kalman update and the associated error covariance as

\[
\begin{align*}
g[n] & = g[n] + \varepsilon [n | n] \\
P_{\varepsilon}[n] & = \delta_{\varepsilon}^n I_{2MNP} \\
\end{align*}
\]

(22)

(23)

where \(\delta_{\varepsilon}^n = \frac{\sigma_{\eta}^2}{2\sigma_{\alpha}^2 + \sigma_{\eta}^2}\).

Assuming the channel estimate at block \(n\) as \(\hat{h}_s[n] = h_s[n | n-1]\) the linear receiver that computes \(\hat{s}_{\alpha,n}\), the estimate of \(s_{\alpha,n}\), is written as

\[
\hat{s}_{\alpha,n} = \frac{1}{||h_s[n]||^2} A^T (h_s[n]) \bar{y}_{n,s} = \bar{s}_i[n] - \bar{v}_{i,n} + \bar{z}_{i,n}
\]

(24)

where the noise term \(\tilde{z}_{i,n}\) is uncorrelated and zero mean with covariance matrix \(R_{\tilde{z}} = \frac{\sigma_{\tilde{z}}^2}{2||h_s[n]||^2} I_{2K}\), and the vector \(\tilde{v}_{i,n}\) zero mean with covariance matrix

\[
R_{\tilde{v}} = \frac{\beta_i^h}{||h_s[n]||^4} A^T (h_s[n]) E_{\tilde{a}_i[n]} \{B(\tilde{s}_i[n]) \\
B^T (\tilde{s}_i[n]) \} A(h_s[n]) 
\]

(25)

where \(E_{\tilde{a}_i[n]}()\) denotes the expectation of the argument with respect to \(x\).

The following lemma derives a property of the OSTBCs, which will be useful to simplify (25).

**Lemma:**
The matrix \(B(\tilde{s}_i[n])\) satisfies \(B(\tilde{s}_i[n]) B^T (\tilde{s}_i[n]) = (\frac{\beta_i^h}{||h_s[n]||^4}) N_{I_{2MT}}\) where \(d\) is the number of zeros contained in each column of \(X(s[n])\).

**proof:** See [10, Appendix B].

Using the above Lemma, (25) is simplified as follows

\[
R_{\tilde{v}} = \left(\frac{\beta_i^h}{||h_s[n]||^4}\right) N_{I_{2K}} = \left(\frac{\beta_i^h}{||h_s[n]||^4}\right) N_{I_{2K}} \text{ where } \beta_i^h = \beta_i^h \left(\frac{\sigma_{\tilde{z}}^2}{||h_s[n]||^4}\right). \]

Now, the SNR of the decision variable can now be written as \(SNR = \frac{\sigma_{\tilde{z}}^2 ||h_s[n]||^2}{2\rho_{\eta}^2 + \sigma_{\eta}^2} \).

Considering \(m\)-ary square QAM, the probability of error can now be approximately written as \([11]\)

\[
P_{e,n} = \frac{2(\sqrt{m-1})}{m \log_2 m} q \left(\sqrt{SNR (m-1)}\right) = c Q(a \sqrt{\psi})
\]

(26)

where \(c = \frac{2(\sqrt{m-1})}{m \log_2 m}, a_n = \frac{\sqrt{\psi}}{\sqrt{1 + \frac{\sigma_{\eta}^2}{\sigma_{\alpha}^2}}} = \frac{||h_s[n]||^2}{\rho_{\eta}}, \text{ and } Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-t^2/2} dt\).

The averaged BER is then written as

\[
\bar{P}_{e,n} = c E_Q(\bar{Q}(a \sqrt{\psi})) = c \int_{0}^{\infty} Q(a \sqrt{\psi}) p_{\xi}(\psi) d\psi
\]

(27)

The integration (27) can be analyzed under two different cases.

**Case I:** \(P = 1\)

In this case we have \(h_s[n] = I_{2MN} g[n]\) where \(g[n]\) is i.i.d Gaussian with zero mean and variance \(\rho_{\eta}\). The probability distribution of \(\psi = ||g||^2 / \rho_{\eta}\) is chi-squared with \(k = 2MN\) degrees of freedom, i.e., \(p_{\xi}(\psi) = \frac{1}{\psi^{k/2} \Gamma(k/2)} \psi^{k/2-1} e^{-\psi/k} \) the moment generating function of the above pdf is given as \((1 - 2k)^{-k/2}\). From [12, pp. 124], the integration of (27) reduces to \(\bar{P}_{e,n} = c \frac{1}{2} \int_0^{\sigma_{\eta}^2} T^{\frac{k}{2}} \left(1 + \frac{\sigma_{\eta}^2}{\sigma_{\alpha}^2} + \frac{\sigma_{\eta}^2}{\sigma_{\alpha}^2}\right)^{m} d\theta \text{ where } \theta = (\frac{\sigma_{\eta}^2}{\sigma_{\alpha}^2})^m\).

Finally, the average bit error rate is written as \(\bar{P}_{e,n} = c \frac{1}{2} \left(1 - \mu_n \sum_{k=0}^{\infty} \left(\frac{2k}{k+1}\right) \left(1 + \frac{\sigma_{\eta}^2}{\sigma_{\alpha}^2}\right)^{m} \right) \text{ where } \mu_n = \sqrt{\frac{\sigma_{\eta}^2}{\sigma_{\alpha}^2}}\).

**Case II:** \(P > 1\)

In this case, the elements of \(h_s[n] = W_s g[n]\) are correlated and hence the distribution of \(\psi\) is complicated and finding a closed form solution for (27) is very challenging. Luckily, the samples of \(\psi\) can be easily generated and hence the following monte carlo integration can be applied to find the averaged BER \(\bar{P}_{e,n} = \frac{1}{N_{s}} \sum_{s=1}^{N_{s}} Q(\psi_s)\) where \(\psi_s \sim p(\psi)\) and \(N_{s}\) is the number of samples taken in the integration.

Now, for both of the above cases, assuming TRP = \(L_s\), the average probability over the \(L_s\) blocks without pilot symbols is obtained as \(\bar{P}_{e} = \frac{1}{L} \sum_{l=1}^{L_s} \bar{P}_{e,n+l-1}\).
signal to noise ratio (SNR) is defined as $SNR = \frac{P}{\sigma_n^2}$ where $P$ is the received signal power at each receiving antenna and $\sigma_n^2$ is the variance of the noise. We consider the Alamouti’s code of [15] for the system with two receiving antennas $(N = T = K = M = 2)$ for the simulation. We study the effect of some of the parameters in the SER performance of the system. Specifically, we will study the effect of the number of continuous blocks without pilot symbols, $N$, number of subcarriers $L$, and $fT$ on the SER performance of the system.

**Example 1 – Effect of the number of sub-carriers**: In Fig. 1 the theoretical study of this example is confirmed through simulation. The figure also confirms that the simulated curves get closer to the theoretical ones as the number of sub-carrier $L$ increases. **Example 2 – Effect $fT$**: In Fig. 2 the simulation is repeated to compare the theoretical curves of Example 3. Here it is intuitive to note that the simulated curves significantly differ from the theoretical ones for higher values of $fT$ due to the higher error in decoding when $fT$ is high.

5. CONCLUSIONS

A theoretical performance analysis scheme is developed to study the bit error rate performance of the Kalman filter based blind channel tracking and data decoding scheme for MIMO-OFDM systems that employ orthogonal space-time codes as the underlying space-time codes. The newly developed schemes enables one to study the effect interference for the system due to different scenarios.

6. REFERENCES


