Propagation channel, capacity and error probability for dual-polarized wireless transmissions

Claude Oestges
Microwave Laboratory, 3 place du Levant, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium
Email: oestges@emic.ucl.ac.be

Received 19 June 2007; accepted 21 August 2007

Use of dual-polarized arrays in multiple-input multiple-output (MIMO) systems is addressed in this paper. A new and simple analytical model of dual-polarized Rayleigh and Ricean fading channels is taken, which relies on a limited number of physical parameters, such as the spatial correlations, the co-polar gain imbalance and the cross-polar discrimination. Then, the multiplexing advantage of dual-polarized transmissions is investigated through the evaluation of the ergodic mutual information. Finally, the performance of two space-time coding schemes (Alamouti O-STBC and un-coded Spatial Multiplexing) is evaluated via a detailed analysis of the pair-wise error probability (PEP).

Keywords: Multiple-input multiple-output (MIMO), Dual-polarized transmission, Space-time coding

PACS No.: 84.40.Ba

1 Introduction

In recent years, increasing attention has been paid to multiple-input multiple-output (MIMO) broadband wireless communication systems. However, sufficiently large antenna spacings are usually required for achieving significant MIMO gains. Hence, the use of co-located orthogonally-polarized antennas appears as a space- and cost-effective alternative\(^1\), as orthogonal polarizations ideally offer a complete separation between channels, with a full decorrelation at both transmit and receive sides. In this paper, the focus is on analysis of 2×2 dual-polarized MIMO channels, i.e. both transmit and receive arrays are made of two antennas with orthogonal polarizations. These antennas may be spatially separated, in order to combine both spatial and polarization diversity/multiplexing. Note that when two co-located antennas with orthogonal polarizations are used at both ends of the link, the channel is defined as a 2×2 system.

Despite a number of recent studies focusing on spatial channel models, only a limited number of papers have addressed the polarization issue\(^1,5\), theoretically or experimentally, mainly because the (de-)coupling effect between orthogonal polarizations is a complex mechanism. Also, a number of studies comparing uni-polarized with dual-polarized communications propose hard conclusions and does not detail under which scenarios polarization may help. Hence, there is a need to clearly identify the benefits of multiple polarizations in the context of simple analytical models, using a reduced number of physically sound parameters.

In Section 2, a simple model is proposed, combining the effects of space and polarization separations. In Section 3, it is investigated whether dual-polarized systems may increase the system capacity. Finally, in Section 4, the performance of Spatial Multiplexing and Orthogonal Space-Time Block Coding (O-STBC) is analyzed from an error probability viewpoint.

2 Channel models for dual-polarized systems

2.1 Antenna and scattering depolarization

Ideally, the cross-polar transmissions (e.g. from a vertically-polarized Tx antenna to an horizontally-polarized Rx antenna) should be equal to zero. This is actually not the case owing to two depolarization mechanisms: antenna depolarization and scattering-induced depolarization. The first mechanism is well-known in antenna theory and is easily accounted for by means of the cross-polar antenna pattern. Analytically, this can be approximated by a scalar antenna cross-polar discrimination \(\chi_a\) (antenna XPD). The latter is used to build an antenna depolarization matrix, which multiplies the channel matrix \(H\), at the transmitter and/or receiver, given by
\[ X_a = \begin{bmatrix} 1 & \sqrt{\chi_a} \\ \sqrt{\chi_a} & 1 \end{bmatrix} \] … (1)

Note that antenna XPD is the only mechanism affecting line-of-sight components. Scattered components (assumed as non-coherent) are affected by both antenna and scattering XPD, although these are well separated effects if the antenna cross-coupling is represented by \( X_a \). For simplicity, it is assumed in what follows that antenna XPD only affects the receiver. In a Ricean channel, the global dual-polarized channel matrix (assuming that the K-factor is identical on all elements) therefore reads as a \( 2 \times 2 \) matrix given by

\[ H_{x,u} = X_a H_x = \sqrt{\frac{K}{K+1}} X_a \tilde{H}_x + \sqrt{\frac{1}{K+1}} X_a \tilde{H}_x \] … (2)

where \( \tilde{H}_x \) is the Ricean coherent (non-fading) \( 2 \times 2 \) matrix, \( \tilde{H}_x \) the Rayleigh non-coherent \( 2 \times 2 \) matrix and \( K \) the co-polar Ricean K-factor, i.e. the ratio of the coherent to the non-coherent part.

### 2.2 Dual-polarized Rayleigh fading channels

In Rayleigh fading, the channel model should account for three mechanisms:

(i) the spatial correlation arising from the finite spacing between the antennas (if dual-polarized antennas are co-located, this correlation is equal to one),

(ii) the gain imbalance between the various co- and cross-polar components, and

(iii) the (de)correlation between all pairs of co- and cross-polar antennas arising only from the polarization difference (i.e. for co-located dual-polarized antennas).

In a first approach\(^6\), \( \tilde{f}_x \) is decomposed by extracting the impact of depolarization on the channel gains, yielding

\[ \tilde{f}_x = \tilde{f}_x |X| \] … (3)

where \( |X| \) depends on the polarization scheme. What is important to notice is that \( \tilde{f}_x \) still includes two correlation mechanisms (space and polarization). Hence, it is generally not equal to an equivalent unipolarized transmission matrix \( \tilde{f}_x \) (i.e. with the same antenna spacings, all polarizations being then identical). As a result, \( \tilde{f}_x \) is some hybrid matrix, modeling the correlation aspects of both spacing and polarization. How does the covariance of \( \tilde{f}_x \) depends on the spatial and polarization correlations is not known a priori, except the use of orthogonal polarizations totally de-correlates all individual channels: \( \tilde{f}_x = H_w \), where \( H_w \) is the classical i.i.d. complex Gaussian matrix, irrespective of the antenna spacing.

In a more general modeling approach\(^7\), a second model explicitly separates spacing-related and polarization-related effects. The separation is thus operated based on the physical mechanisms (space versus polarization) rather than on their impact (gain versus correlation). Subsequently, the dual-polarized Rayleigh channel matrix may be rewritten as

\[ \tilde{f}_x = \tilde{f}_0 X \] … (4)

In Eq. (4), \( \tilde{f}_0 \) is modeled as a uni-polarized correlated Rayleigh channel, while \( X \) models both the correlation and power imbalance impacts of scattering-induced depolarization. It is important to stress that \( X \) only models the power imbalance and the phase-shifts between the four channels, but does not contain fading. Normalized fading (i.e. with unit average power) is entirely modeled by \( \tilde{f}_0 \), with

\[ \text{vec}(\tilde{f}_0) = R^{1/2} \text{vec}(H_w) \] … (5)

where vec is the operator stacking all the elements of the \( 2 \times 2 \) matrix columnwise into a \( 4 \times 1 \) vector, and \( H_w \) is one realization of the spatially uncorrelated \( 2 \times 2 \) Rayleigh channel. In Eq. (5), \( R \) is the \( 4 \times 4 \) uni-polarized correlation matrix of \( \text{vec}(\tilde{f}_0) \), given by

\[ R = \begin{bmatrix} 1 & t & r & s_1^* \\ t & 1 & s_2^* & r^* \\ r & s_2 & 1 & t^* \\ s_1 & r & t & 1 \end{bmatrix} \] … (6)

where \( r \) and \( t \) are the usual receive and transmit correlation coefficients, and \( s_1 = E\{h_{11}, h_{22}^*\} \) and \( s_2 = E\{h_{12}, h_{21}^*\} \) known as the cross-channel correlation coefficients\(^8\). A relatively general model of the channel matrix for VH-to-VH downlink transmission\(^7\) is given similarly to Eq. (5), given by
where \( \mu \) and \( \chi \) represent respectively, the co-polar imbalance and the scattering XPD, and are assumed to be constant. Also, \( \sigma \) and \( \vartheta \) are receive and transmit correlation coefficients (i.e. the correlation coefficients between VV and HV, HH and HV, VV and VH or HH and VH for co-located antennas), while \( \delta_1 \) and \( \delta_2 \) the cross-channel correlation coefficients caused by the use of orthogonal polarizations. That is \( \delta_1 \) is the correlation between the VV and the HH components, and \( \delta_2 \) the correlation between the VH and the HV components. Again \( \mathbf{X}_w \) is a 2 \( \times \) 2 matrix whose four elements are independent circularly symmetric complex exponentials of unit amplitude, \( \exp(\pm j\varphi_k), \ k = 1, \ldots, 4 \), the angles \( \varphi_k \) being uniformly distributed over (0, 2 \( \pi \)).

In the following, it is assumed for simplicity that the correlation coefficients \( \sigma \) and \( \vartheta \) between any cross-polar component (VH or HV) and any co-polar component (VV or HH) are equal to zero, although they might actually be slightly higher\(^9\).

### 2.3 Dual-polarized Ricean fading channels

In Ricean fading channels, the Ricean part of the dual-polarized matrix is expressed (without the antenna XPD effect) for an arbitrary 2 \( \times \) 2 symmetric scheme as

\[
\mathbf{H}_x \approx \begin{pmatrix}
1 & \sqrt{\chi} \\
\sqrt{\mu} & \mu \\
\sqrt{\mu} & \sqrt{\mu} \\
\end{pmatrix} \mathbf{H}_w \mathbf{X}_w
\]

Now one can consider two particular cases\(^10\), i.e. \( \chi_a = 0 \) and \( \chi_a = 1 \). The mutual information is then given by

\[
I(\chi_a =0) = \log_2 \left(1 + \rho/2\right)
\]

\[
I(\chi_a =1) = \log_2 \left(1 + 2\rho\right)
\]

Both cases are compared in Fig. 1. It is clear that a good antenna XPD restores the multiplexing gain at high SNR levels. Here, the mutual information of a two-input two-output scheme with equal power allocation across the transmit antennas is given by

\[
I(\mathbf{I}_2) = \log_2 \left(1 + \rho/2\right)
\]

Two scenarios are considered:

(i) Spatially correlated antennas and \( \delta_1 = \delta_2 = 1 \), where

\[
\mathbf{H}_r = \mathbf{I}_2, \quad \mathbf{H}_w = \begin{pmatrix}
1 & \sqrt{\chi} e^{j\varangle} \\
\sqrt{\mu} e^{j\varangle} & \mu \\
\end{pmatrix}
\]

### 3 Mutual information of dual-polarized channels

For very high K-factors, the channel matrix in Eq. (2) only depends on the Ricean component. Assuming that the latter is dominated by the line-of-sight contribution, the dual-polarized channel matrix in a broadside array configuration reads as

\[
I_I = \frac{1}{2} \log \left(1 + \rho/2\right)
\]

\[
I_I = \frac{1}{2} \log \left(1 + 2\rho\right)
\]

Now a finite K-factor is assumed, and the model of Eq. (2) used to highlight the impact of scattering XPD and gain imbalance on the mutual information upper bound. For convenience, one can also take \( \chi_a = 0 \). Two scenarios are considered:

(i) Spatially correlated antennas and \( \delta_1 = \delta_2 = 1 \), where

\[
\mathbf{H}_r = \mathbf{I}_2, \quad \mathbf{H}_w = \begin{pmatrix}
1 & \sqrt{\chi} e^{j\varangle} \\
\sqrt{\mu} e^{j\varangle} & \mu \\
\end{pmatrix}
\]

where \( \mathbf{H}_r \) is given by Eq. (5),

![Fig. 1—Mutual information of uni- and dual-polarized 2 \( \times \) 2 Ricean fading channels](image_url)
(ii) Well separated (un-correlated) antennas, where

\[
\mathbf{H}_x = \mathbf{1}_2, \quad \mathbf{H} = \left[ \begin{array}{c}
\frac{1}{\sqrt{\chi}} \\
\frac{\sqrt{\mu}}{\sqrt{\chi}} \end{array} \right] \mathbf{H}_x
\]

which is valid irrespective of \(\delta_1\) and \(\delta_2\).

Two alternative scenarios (i.e. \(\delta_1 = \delta_2 = 0\) with close or well separated antennas) do not need being considered, as they are indeed both covered by the second scenario (all three scenarios cause elements of \(\mathbf{H}_x\) to become fully un-correlated). For both analyzed scenarios, one can resort to an upper bound\(^{10}\) of the ergodic (i.e. averaged over the fading distribution) mutual information \(I\), outlined by

\[
E\{I\} \leq \log_2 \left( \frac{1}{\mu} \left[ \begin{array}{c}
1 + \frac{\rho}{\sqrt{\mu}} \mathbf{H} \mathbf{H}^H \\
\end{array} \right] \right)
\]

where \(E\) designates the expectation operator. In the considered cases, the upper bounds become\(^{11}\)

(i) For \(\delta_1 = \delta_2 = 1\),

\[
\mathbf{r} = 1 + \frac{\rho^2}{2} K + (1 + \mu)(1 + \chi) + \left( \frac{\rho}{\sqrt{\mu}} \mathbf{H} \mathbf{H}^H \right)^2
\]

\[
\times \left( \begin{array}{c}
1 \frac{K}{(K+1)} \left[ \left( \begin{array}{c}
1 + \frac{\mu}{K+1} \frac{\mathbf{H} \mathbf{H}^H}{\sqrt{\mu}} \end{array} \right) \right] \\
\end{array} \right)
\]

(ii) For well-separated dual-polarized antennas with arbitrary values of \(\delta_1\) and \(\delta_2\) or for \(\delta_1 = \delta_2 = 0\),

\[
\mathbf{r} = 1 + \frac{\rho^2}{2} K + (1 + \mu)(1 + \chi) + \left( \frac{\rho}{\sqrt{\mu}} \mathbf{H} \mathbf{H}^H \right)^2
\]

\[
\times \left( \begin{array}{c}
1 \frac{K}{(K+1)} \left[ \left( \begin{array}{c}
1 + \frac{\mu}{K+1} \frac{\mathbf{H} \mathbf{H}^H}{\sqrt{\mu}} \end{array} \right) \right] \\
\end{array} \right)
\]

It is first observed that if \(K \rightarrow \infty\), both Eqs (14) and (15) approach the result of the first case in Eq. (10). Then it is also noted that the mutual information is higher in the first scenario (i.e. \(\delta_1 = \delta_2 = 1\)). Indeed, in the latter, the channel matrix behaves as a diagonal channel (for \(\mu = \chi = 1\) and \(K = 0\), \(\mathbf{H}_x = \mathbf{H}^H\) is exactly a \(2 \times 2\) diagonal channel\(^{12}\)). Note however that the difference between achieved mutual information values in both scenarios remains limited. Finally, the ergodic mutual information of Rayleigh channels increases in both Eqs (14) and (15) as \(\mu\) and \(\chi\) increase, as illustrated in Figure 2 for \(\chi\) (assuming that \(\mu = 0.5\) and close antennas with \(\delta_1 = \delta_2 = 1\)). It is also observed that dual-polarized transmissions only offer larger mutual information for Ricean or highly correlated Rayleigh fading channels.

4 Average PEP in dual-polarized channels

It is now of interest to find the average PEP as a function of the SNR \(\rho\), i.e. the probability that the receiver decodes the codeword \(\mathbf{E} = [c_0, \mathbf{L}, e_{T-1}]\) instead of codeword \(\mathbf{C} = [c_0, \mathbf{L}, e_{T-1}]\) (\(T\) is the duration of the space-time block code). In \(2 \times 2\) dual-polarized Rayleigh and LOS Ricean slow fading channels, the approximate expressions of the PEP\(^{7}\) are respectively given

\[
P(C \rightarrow E) \leq \prod_{k=1}^{T} \left( 1 + \frac{\rho}{4 \chi} \mathbf{C} \mathbf{R} \mathbf{C}^H \right)
\]

and

\[
P(C \rightarrow E) = Q \left[ \sqrt{\frac{\rho}{2K+1}} \left( \mathbf{H} \mathbf{(C-E)(C-E)^H} \right) \right]
\]

where \(\mathbf{C} = \mathbf{R}_{n \times n} \left[ \mathbf{I} \otimes (\mathbf{E}) \mathbf{(C-E)^H} \right] \). \(\mathbf{R}_{n \times n}\) being the correlation matrix of \(\text{vec}(\mathbf{H}^H)\).

Fig. 2—Mutual information of uni- and dual-polarized \(2 \times 2\) channels for different \(K\)-factors and correlations
In this paper, the analysis is kept restricted to two simple schemes:

(i) the Alamouti O-STBC \((T = 2)\), which extends the principle of transmit/receive diversity,

(ii) the Spatial Multiplexing scheme, which consists in sending different data streams over each antenna, thereby increasing the system throughput.

In Rayleigh fading channels with vertical and horizontal antennas at both ends, neglecting antenna XPD \((\chi_a = 0)\) and using the same models as those used when estimating the mutual information, one may write the dual-polarized correlation matrix as

\[
R_{\chi,a} = \begin{bmatrix}
1 & 0 & 0 & \sqrt{\mu_s} \\
0 & \mu \chi & \sqrt{\mu \chi^2} & 0 \\
0 & \sqrt{\mu \chi^2} & \chi & 0 \\
\sqrt{\mu_s} & 0 & 0 & \mu \\
\end{bmatrix}
\]  \hspace{1cm} (17-a)

if \(\delta_1 = \delta_2 = 1\), and

\[
R_{\chi,a} = \text{diag}\{1, \mu \chi, \chi, \mu\}
\]  \hspace{1cm} (17-b)

if \(\delta_1 = \delta_2 = 0\).

4.1 Performance of orthogonal space-time block coding

Restricting the present analysis to the Alamouti scheme, the codeword matrix is given by

\[
C = \begin{bmatrix}
c_0 & -c_1^* \\
c_1 & c_0^* \\
\end{bmatrix}
\]  \hspace{1cm} (18)

where \(c_0\) and \(c_1\) are two symbols of a given constellation, and the codeword errors are denoted as

\[
d_0 = \left|c_0 - e_1\right|/\sqrt{2} \quad \text{and} \quad d_1 = \left|c_1 - e_1\right|/\sqrt{2}.
\]

In Rayleigh fading channels, the matrix \(C_{R_{\chi,a}}\) is rewritten for O-STBC as

\[
C_{R_{\chi,a}} = R_{\chi,a} \left[|d_0|^2 + |d_1|^2\right]
\]  \hspace{1cm} (19)

with \(R_{\chi,a}\) being given by Eqs (6), (17-a) or (17-b).

Therefore, the impact of the propagation channel on the error probability is directly given by the four eigenvalues of the correlation matrix. Minimizing Eq. (16-a) comes to maximize \(\det[I_z + \rho|\text{R}|]\), where an effective SNR is defined for better legibility as

\[
\rho_z = \rho \left(|d_0|^2 + |d_1|^2\right)/4.
\]

Dual-polarization is then preferred when

\[
\det[I_z + \rho|\text{R}|] > \det[I_z + \rho|\text{R}|] \quad \text{... (20)}
\]

In the high SNR regime, the condition simply becomes

\[
\det R_x > \det \text{R} \quad \text{... (21)}
\]

Note however that the high SNR regime might only be reached for unrealistically high SNR when \(\chi\) and/or \(\mu\) are small, i.e. when some eigenvalues of \(R_x\) are small. The above conditions are easily expressed in terms of the correlation coefficients, as well as \(\chi\) and \(\mu\). As an example, consider a Kronecker-structured channel un-correlated at the receiver (hence, \(r = s_1 = s_2 = 0\)). When \(\mu = 1\) (i.e. when co-polarized waves are equally attenuated), the above condition becomes

\[
|r| > \sqrt{1 - \chi^2}.
\]

A similar condition on \(|r|\) would be found for Kronecker-structured channels un-correlated at the transmitter. This indicates that dual-polarization can increase the performance at high spatial transmit/receive correlation levels, despite the reduction of the average energy of the channel caused by \(\chi\). Naturally, the higher the value of \(\chi\), the lower the transmit/receive correlation for which dual-polarized schemes outperform uni-polarized schemes. Note also that transmit and receive correlations equally affect the performance of diversity schemes.

To analyze Ricean fading channels, one can assume that the Ricean component is the line-of-sight, and that the K-factor is high (so that the Rayleigh component can be neglected). Then, the performance solely depends on the Frobenius norm of the coherent component. Naturally, this norm is always smaller for dual-polarized schemes (since \(1 < \chi\)). Therefore, the use of dual-polarized transmissions is not recommended for diversity schemes in Ricean channels.

4.2 Performance of spatial multiplexing

For a simple un-coded spatial multiplexing scheme, one can have
In Rayleigh fading, Eq. (22) implies that the rank of both $C_R$ and $C_{R_c}$ is two, with their two non-zero eigenvalues $\lambda_{1,2}$ given by

$$\lambda_{1,2} = \frac{a + f \pm \sqrt{(a-f)^2 + 4b^2}}{2}$$  \hspace{1cm} (23)

with

(i) For uni-polarized schemes

\begin{align*}
a &= |d_0|^2 + |d_1|^2 + 2\Re\{d_0 d_1^*\} \\
b &= r(|d_0|^2 + |d_1|^2) + s_1 d_0 d_1^* + s_2 d_0^* d_1 \\
f &= a
\end{align*}

(ii) For dual-polarized schemes with $\delta_1 = \delta_2 = 1,$

\begin{align*}
a &= |d_0|^2 + \mu \chi |d_1|^2 \\
b &= \sqrt{\mu} (s_1 d_0 d_1^* + s_2 d_0^* d_1) \\
f &= \chi |d_0|^2 + \mu |d_1|^2
\end{align*}

(iii) For dual-polarized schemes with $\delta_1 = \delta_2 = 0,$

\begin{align*}
a &= |d_0|^2 + \mu \chi |d_1|^2 \\
b &= 0 \\
f &= \chi |d_0|^2 + \mu |d_1|^2
\end{align*}

An estimate of the average symbol error rate may be easily obtained by weighting the average PEP over the different possible symbols, using a union bound approximation. Various simulation results are illustrated in Fig. 3 for a QPSK modulation.

Considering first uni-polarized systems, one can observe that transmit correlation is more harmful than receive correlation \(^{13}\). As for the use of dual-polarized arrays, it is clear that these are only beneficial when transmit and/or receive correlations are higher than a certain level. This level is directly related to the co-polar gain imbalance and the scattering XPD. The smaller the value of $\chi$ and/or $\mu$, the larger this correlation level. Naturally, when both $\chi$ and $\mu$ are equal to one, dual-polarized systems always perform better than uni-polarized systems, as the de-correlation effect is not hampered by a decrease of power. For $\chi$ and/or $\mu$ smaller than one, the required correlation increases rapidly. As a consequence, dual-polarized spatial multiplexing schemes perform well in Rayleigh fading channels when the uni-polarized spatial correlations are high enough.

In LOS Ricean fading channels, the error probability is obtained through Eq. (16-b). Simulation results are displayed in Fig. 4 for $\chi = \mu = 1$ and $\delta_1 = \delta_2 = 0$ at 10 dB SNR. The use of dual-polarized systems significantly increases the performance. Note that the
SER for uni-polarized links is indeed given by $\chi_a = 1$ when $K$ is high enough, so that the Rayleigh component can be neglected.

5 Conclusions

A simple analytical model of dual-polarized $2 \times 2$ MIMO transmissions has been presented, which efficiently separates the spatial and polarization effects. As far as the benefits from dual-polarized arrays are concerned, the derivations point out that large multiplexing gains are achievable in Ricean or highly correlated Rayleigh fading channels, but that diversity gains are only possible in highly correlated Rayleigh fading channels (and not in Ricean channels). The analysis also identifies quantitatively, for which channels (correlations, K-factor) schemes (O-SBTC or spatial multiplexing) and SNR levels dual-polarization schemes may turn beneficial.

References