Design Criterion of Space-Time Codes for SDM Optical Fiber Systems

El-Mehdi Amhoud, Ghaya Rekaya-Ben Othman, Yves Jaouën
LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 46 rue Barrault, 75634 Paris, France
Email: el-mehdi.amhoud@telecom-paristech.fr

Abstract—The last two decades have known an exponential growth in the demand for network bandwidth. Since single mode fiber systems are approaching the nonlinear Shannon limit, multimode fibers (MMF) holds the promise to increase the capacity of optical transmission systems. Propagating modes through multimode fibers are affected by a non-unitary crosstalk known as mode dependent loss (MDL). The impact of MDL and its detrimental effect on the channel capacity was reported in many studies. Space-time coding already designed for wireless communications, was proved to mitigate the non-unitary effects of mode dependent loss. In this paper we derive an upper bound of the error probability of the optical channel affected by MDL, this upper bound yields the design criterion of space-time codes allowing total mitigation of MDL in space division multiplexed transmission systems.

Index Terms—Mode Dependent Loss, Space-Time Coding, Error Probability, Optical Fiber Communication, MIMO

I. INTRODUCTION

Optical transmission systems are fundamental for telecommunication infrastructure. For long-haul as well as metropolitan and access networks, this technology can carry data over long distances with low attenuation. Originally, the first optical links were non-coherent and the modulation format used is the so called On-Off keying. This technique has a low cost and requires only a low amount of optical components but uses only the amplitude of the electromagnetic field. To increase the capacity of optical fiber links coherent detection was introduced. This technology allows the detection of the amplitude and the phase of signals and hence the use of higher order modulation formats. In addition to coherent detection, wavelength division multiplexing (WDM) and advancement in optical amplifiers enabled the transmission of independent modulated wavelengths without opto-electrical regeneration [1]. Moreover, polarization division multiplexing (PDM) can double the capacity of a link by using the two orthogonal polarizations of the electromagnetic field. However WDM-PDM systems are approaching their nonlinear limit. This nonlinear property is due to the increasing amount of light injected in a small volume [2], [3].

The last twenty years have known an exponential growth in the demand for more network capacity, this growth was mainly caused by the built out of the Internet and the growing traffic generated by an increasing number of users. Since frequency, time, phase, polarization have already been used to satisfy the demand for bandwidth, space [4] remains the only available degree of freedom that can be used in optical transmission systems in order to increase the capacity. Space division multiplexing (SDM) can be realized through multimode fibers that allow the propagation of many modes in a single core or multi-core fibers (MCF) where each core can be single-mode or multimode. In the ideal version, SDM can multiply the capacity of a link by the number of propagating modes. In real transmission systems, modes are affected by cross-talk especially if cores are close in MCF or if differential mode group delay (DMGD) is close to zero in MMF. Multiple-input multiple-output (MIMO) decoding techniques already used in wireless communications allow to recover the signals at the receiver. Moreover propagating modes are affected by a non-unitary cross-talk known as mode dependent loss arising from fiber imperfections [5] (splices, microbends) and optical components (amplifiers, multiplexers). The impact of MDL and its detrimental effect on the channel capacity was reported in several studies [4], [6], [7].

In this work, we are interested in SDM system based MMF which can be represented as a $M \times M$ MIMO system, where $M$ modes are emitted (resp. detected) at the transmitter (resp. receiver). Optical solutions such as mode scrambling or strong mode coupling fibers were suggested to reduce the impact of MDL on the channel capacity. In [8] the authors show that placing mode scramblers between fiber spans mitigates MDL, however this technique needs a high number of scramblers which does not introduce any extra MDL. In a recent work [9], a digital signal processing technique based on space-time coding was proved to mitigate MDL in mode division multiplexed systems. For a $6 \times 6$ MIMO SDM system and a link MDL of $10$ dB the authors demonstrated that space-time codes originally designed for wireless MIMO systems completely absorbs the SNR penalty. In this paper, we investigate through theoretical analysis the benefit of space-time coding for the SDM optical channel affected by MDL. We derive a design criterion for space-time codes construction based on the error probability upper bound, we also analyze the behavior of some existing codes and compare their performance.

This paper is organized as follows: In Section II we start by describing a mathematical channel model for SDM optical channel affected by MDL and review the effect of MDL on the system capacity. In Section III we derive an upper bound of the error probability. In Section IV we define a design criterion of codes that minimizes the error probability of a transmission system through multimode fibers affected by...
MDL. In Section V we investigate the performance of some existing space-time codes. Finally in Section VI we conclude and set forth the perspectives of our work.

II. CHANNEL MODEL

A. Optical Channel with MDL

The effect of mode dependent loss on propagation has been investigated in many studies [4], [6], [7]. The link MDL depends on the number and nature of concatenated fiber spans. In our work we consider a mathematical channel model made of a single lumped MDL element to derive the error probability upper bound. This model was used in [10] to study the bit error rate performance of the optical channel affected by MDL. To focus on the impact of MDL on system performance, polarization crosstalk is not considered and we neglect fiber nonlinearities. Differential mode group delay is also not considered since it does not impact the system capacity and can be managed using OFDM format with a cyclic prefix larger than the maximum DMG. The multiple-input multiple-output system may be described by:

\[ Y_{M \times T} = H X_{M \times T} + N_{M \times T} \]  

(1)

\( X_{M \times T} \) is a space-time codeword, where at each instant a linear combination of information symbols is sent on each mode. \( Y_{M \times T} \) is the received codeword where \( T \) represents the temporal codelength. \( N_{M \times T} \) is the noise assumed to be additive white Gaussian with variance \( 2\sigma^2 \) per complex dimension. The channel matrix is given by:

\[ H = \sqrt{\alpha} \sqrt{D} U. \]  

(2)

\( D \) is a diagonal matrix, its entries are uniformly drawn in \([\lambda_{\min}, \lambda_{\max}]\) and represent different losses experienced by each mode. In the absence of MDL the matrix \( D \) is equal to the identity matrix (\( D = I \)). \( U \) is a unitary matrix, it represents coupling between modes. \( \alpha = \frac{M}{\sum_{i=1}^{M} \lambda_i} \) factors out the mode average propagation loss such that \( \text{Trace}(HH^*) = M \).

B. Mode Dependent Loss

There are two different definitions of mode dependent loss in literature. In [11] it is defined as the standard deviation (std) of the channel matrix eigenvalues as follows: MDL = std(\( \lambda_i \)). In [4], [6], [7] mode dependent loss is defined as the ratio of the maximum to the minimum eigenvalues of the channel matrix:

\[ \text{MDL}_{(dB)} = 10 \log(\text{MDL}) = 10 \log(\frac{\lambda_{\max}}{\lambda_{\min}}). \] 

We adopt this last definition in the rest of our work.

The capacity of a MIMO-SDM system is given by:

\[ C = \sum_{i=1}^{M} \log_2(1 + \frac{P}{M} \lambda_i) \]  

(3)

Where \( P \) is the signal to noise ration. In the absence of MDL, space division multiplexing based multimode fiber multiplies the capacity of a link by the number of propagating modes, in this case all modes have the same attenuation (\( \lambda_i = 1 \)). Due to MDL this capacity is reduced and in the worst case the fiber supports the propagation of only one spatial mode [11].

III. AN UPPER BOUND OF THE ERROR PROBABILITY

In the following we suppose that \( H \) is known at the receiver, the error probability is defined as:

\[ P_{\text{error}} = \sum_{X_i \in C} P_r(X_i) \Pr(X_j \neq X_i/X_i) \]  

(4)

Where \( X_i \) (resp. \( X_j \)) is the transmitted (resp. the estimated) codeword. For equiprobable codewords the error probability can be upper bounded by [13]:

\[ P_{\text{error}} \leq \sum_{X_i \in C} \frac{1}{\text{card}(C)} \sum_{X_i \neq X_j} P_r(X_i \rightarrow X_j) \]  

(5)

Where \( C \) is the codebook made of all possible codewords and \( \text{card}(C) \) is the cardinality of \( C \). \( P_r(X_i \rightarrow X_j) \) is the pairwise error probability given by:

\[ P_r(X_i \rightarrow X_j) = Q(\frac{\|H(X_i - X_j)\|^2}{2\sigma}) \]  

(6)

Using Chernoff’s bound and averaging over channel realizations we get:

\[ P_r(X_i \rightarrow X_j) \leq \mathbb{E}[\text{exp}(-\frac{\|H(X_i - X_j)\|^2}{8\sigma^2})] \]  

(7)

Let the minimum euclidean distance of the codebook be: \( d_{\min}^2 = \min_{X_i \neq X_j} \|X_i - X_j\|^2 \). The number of nearest neighbors \( N_{\min,i} \) of \( X_i \) is defined as the number of codewords that are at \( d_{\min} \) of \( X_i \).

Let \( X = X_i - X_j \) denote the difference of two codewords, we obtain:

\[ P_r(X_i \rightarrow X_j) \leq \mathbb{E}[\text{exp}(-\frac{\|\sqrt{\sum_{i=1}^{M} \lambda_i} \sqrt{D} U X\|^2}{8\sigma^2})] \]  

\[ \leq \mathbb{E}[\text{exp}(-\frac{M}{\sum_{i=1}^{M} \lambda_i} \text{Tr}(DUX^*U^*)}] \]  

(8)

In the coming we derive the error probability upper bound for orthogonal space-time codes, afterwards we deal with the general case of space-time codes.

A. Orthogonal Space-Time Codes

Orthogonal space-time codes were widely studied for wireless communication. Thanks to code orthogonality, the maximum-likelihood decoding of these codes is equivalent to a linear processing which reduces the decoding complexity [14].

The main drawbacks of these code family is the transmission rate limitation. For example, the famous Alamouti code has a rate of \( \frac{1}{2} \) sym/c.u (symbol/channel use) when used for a \( 2 \times 2 \) MIMO system. Moreover, construction properties of these codes make the codeword difference matrix inherits the orthogonality property of the code [15]: \( XX^* = \sum_{k} [x_{k,i} - x_{k,j}]^2 \mathbb{I} \) for all possible codeword differences, with \( x_{k,i} \) (resp. \( x_{k,j} \)) the
emitted (resp. estimated) symbols. Applying this property to equation (9) we obtain:

$$P_r(X_i \rightarrow X_j) \leq \exp(-\frac{\|X\|^2}{8\sigma^2}) \tag{10}$$

Reporting this in equation (5) we obtain:

$$P_e \leq \sum_{X_i \in C} \frac{1}{\text{card}(C)} \sum_{X_i \neq X_j} \exp(-\frac{\|X\|^2}{8\sigma^2}) \tag{11}$$

By using the number of nearest neighbors \(N_{\text{min},i}\) of \(X_i\) we obtain:

$$P_e \leq \left(\frac{1}{\text{card}(C)}\right) \sum_{X_i \in C} N_{\text{min},i} \cdot \exp(-\frac{d^2_{\text{min}}}{8\sigma^2}) \tag{12}$$

We denote \(\overline{N}_{\text{min}} = \frac{1}{\text{card}(C)} \sum_{X_i \in C} N_{\text{min},i}\) the average number of closest neighbors of \(X_i\), equation (12) becomes:

$$P_e \leq \overline{N}_{\text{min}} \cdot \exp(-\frac{d^2_{\text{min}}}{8\sigma^2}) \tag{13}$$

From equation (13), we notice that orthogonal space-time codes gives an upper bound completely independent of MDL.

**B. General Case of Space-Time Codes**

In this section we derive an upper bound of the error probability in the general case of space-time codes. We rewrite inequality (5) as a summation of two terms where the first one contains the orthogonal codeword differences and the second term contains the non orthogonal codeword differences.

$$P_e \leq \sum_{X_i \in C} \frac{1}{\text{card}(C)} \sum_{X_i \neq X_j} P_r(X_i \rightarrow X_j) \tag{14}$$

The first term of equation (14) was computed in the previous section and is equal to \(\overline{N}_{\text{min}} \cdot \exp(-\frac{d^2_{\text{min}}}{8\sigma^2})\), where \(\overline{N}_{\text{min}}\) is the average number of nearest neighbors of \(X_i\) such that \(X\) is orthogonal. In the coming, we compute \(P_r(X_i \rightarrow X_j)\) such that \(X\) is non orthogonal. Starting from equation (9) and since \(\lambda_i \leq M\lambda_{\text{max}}\) we get:

$$P_r(X_i \rightarrow X_j) \leq \mathbb{E}_{D,U}[\exp(-\frac{T_r(DUX^*U^*)}{8\sigma^2\lambda_{\text{max}}})] \tag{15}$$

Here, we average over the diagonal and unitary matrix entries. \(XX^*\) is a square hermitian matrix, so there exists a unitary matrix \(V\) and a diagonal matrix \(\Sigma = \text{diag}(\sigma_1, ..., \sigma_M)\) such that: \(XX^* = V \Sigma V^*\). We obtain:

$$P_r(X_i \rightarrow X_j) \leq \mathbb{E}_{D,U}[\exp(-\frac{T_r(DUV\Sigma V^*U^*)}{8\sigma^2\lambda_{\text{max}}})] \tag{16}$$

The matrix \(U\) is randomly drawn from the unitary matrices ensemble, it induces that the product \(UV\) follows the same distribution as \(U\) [16]. We obtain:

$$P_r(X_i \rightarrow X_j) \leq \mathbb{E}_{D,U}[\exp(-\frac{T_r(DU\Sigma U^*)}{8\sigma^2\lambda_{\text{max}}})] \tag{17}$$

by developing the term \(DU\Sigma U^*\) we obtain:

$$P(X_i \rightarrow X_j) \leq \mathbb{E}_{D,U}[\exp(-\frac{\sum_{i,j=1}^{M} \lambda_i \sigma_j |u_{ij}|^2}{8\sigma^2\lambda_{\text{max}}})] \tag{18}$$

$$\leq \mathbb{E}_{D,U}[\prod_{i=1}^{M} \exp(-\lambda_i \sum_{j=1}^{M} \sigma_j |u_{ij}|^2)] \tag{19}$$

$$\leq \prod_{i=1}^{M} \mathbb{E}_{U}[\exp(-\lambda_i \sum_{j=1}^{M} \sigma_j |u_{ij}|^2)] \tag{20}$$

We average over \(\lambda_i\)’s that are independently uniformly drawn in \([\lambda_{\text{min}}, \lambda_{\text{max}}]\).

$$P(X_i \rightarrow X_j) \leq \prod_{i=1}^{M} \mathbb{E}_U \left[ \exp(-\frac{\sum_{j=1}^{M} \lambda_{\text{min}} \sigma_j |u_{ij}|^2}{8\sigma^2\lambda_{\text{max}}}) \times P(\lambda_i) d\lambda_i \right] \tag{21}$$

$$= \prod_{i=1}^{M} \mathbb{E}_U \left[ \frac{\exp(-\sum_{j=1}^{M} \sigma_j |u_{ij}|^2)}{\lambda_{\text{max}} - \lambda_{\text{min}}} \right] \times \frac{\exp(-\sum_{j=1}^{M} \lambda_{\text{min}} \sigma_j |u_{ij}|^2)}{8\sigma^2\lambda_{\text{max}}} \right] \tag{22}$$

$$= \prod_{i=1}^{M} \mathbb{E}_U \left[ \frac{\exp(-\frac{1}{2} \left(1 + \frac{\lambda_{\text{min}}}{\lambda_{\text{max}}} \right) \sum_{j=1}^{M} \sigma_j |u_{ij}|^2)}{8\sigma^2\lambda_{\text{max}}} \right] \times \frac{2 \sinh \left( \frac{1}{2} \left(1 - \frac{\lambda_{\text{min}}}{\lambda_{\text{max}}} \right) \sum_{j=1}^{M} \sigma_j |u_{ij}|^2 \right)}{8\sigma^2\lambda_{\text{max}}} \right] \tag{23}$$

In equation (21) \(P(\lambda_i)\) represents the probability distribution function of \(\lambda_i\) and it is given by:

$$P(\lambda_i) = \begin{cases} \frac{1}{\lambda_{\text{max}} - \lambda_{\text{min}}} & \text{if } \lambda_{\text{min}} \leq \lambda_i \leq \lambda_{\text{max}} \\ 0 & \text{elsewhere} \end{cases}$$

We use the approximation of the hyperbolic sine at high SNR: \(\sinh(x) = \exp(x)/2 \) so we get:

$$P(X_i \rightarrow X_j) \leq \prod_{i=1}^{M} \mathbb{E}_U \left[ \exp\left(-\frac{\lambda_{\text{min}} \sum_{j=1}^{M} \sigma_j |u_{ij}|^2}{8\sigma^2\lambda_{\text{max}}} \right) \right] \tag{24}$$

$$= \mathbb{E}_U \left[ \exp\left(-\frac{\lambda_{\text{min}} \sum_{j=1}^{M} \sigma_j |u_{ij}|^2}{8\sigma^2\lambda_{\text{max}}} \right) \right] \tag{25}$$

$$= \mathbb{E}_U \left[ \exp\left(-\frac{\lambda_{\text{min}} \sum_{j=1}^{M} \sigma_j |u_{ij}|^2}{8\sigma^2\lambda_{\text{max}}} \right) \right] \tag{26}$$

Equation (27) is independent of the unitary matrix \(U\) so we get:

$$P(X_i \rightarrow X_j) \leq \exp\left(-\frac{\|X\|^2}{8\sigma^2\lambda_{\text{max}}} \right) \tag{28}$$
Substituting (28) in the second term of equation (14) gives:
\[
P_e \leq N_{1_{\text{min}}} \cdot \exp\left(\frac{d_{\text{min}}^2}{8\sigma^2}\right) + \sum_{X_i \in C} \frac{1}{\text{card}(C)} \sum_{X_j \neq X_i} \exp\left(-\frac{||X||^2}{8\sigma^2 \text{MDL}}\right) \tag{29}
\]
Let \(N_{2_{\text{min}}}\) be the number of nearest neighbors of \(X_i\) such that \(X\) is non orthogonal, we obtain:
\[
P_e \leq N_{1_{\text{min}}} \cdot \exp\left(\frac{d_{\text{min}}^2}{8\sigma^2}\right) + \sum_{X_i \in C} \frac{1}{\text{card}(C)} N_{2_{\text{min},i}} \exp\left(-\frac{d_{\text{min}}^2}{8\sigma^2 \text{MDL}}\right) \tag{30}
\]
by denoting \(N_{2_{\text{min}}} = \sum_{X_i \in C} \frac{1}{\text{card}(C)} N_{2_{\text{min},i}}\) we arrive at:
\[
P_e \leq N_{1_{\text{min}}} \cdot \exp\left(-\frac{d_{\text{min}}^2}{8\sigma^2}\right) + N_{2_{\text{min}}} \cdot \exp\left(-\frac{d_{\text{min}}^2}{8\sigma^2 \text{MDL}}\right) \tag{31}
\]
From equation (31), the error probability upper bound is composed of two terms. The first one comes from the orthogonal codeword differences, it is completely independent of MDL. The second term is affected by MDL and comes from the non orthogonal difference codewords. In order to minimize the error probability upper bound the first term should be the dominant term of the upper bound.

**Proposition:** In order to minimize the error probability, the average number of nearest neighbors \(N_{1_{\text{min}}}\) of a space-time code such that \(X\) is orthogonal should be maximized.

The design criterion obtained is directly related to the orthogonality of codeword differences, it is completely different from the rank and the determinant criteria of a Rayleigh fading channel. In Table I, we have computed the average number of orthogonal and non-orthogonal closest neighbors of \(X_i\) for different space-time codes. We notice that the Alamouti code has \(N_{2_{\text{min}}} = 0\), this is due to the orthogonal structure of this code. The codewords of the Silver code have more orthogonal nearest neighbors (6.5 neighbors) than the codewords of the TAST code (4 neighbors). Codewords of The golden code have no nearest orthogonal neighbors. This observation makes us think that the silver code gives the best performance followed by the TAST then the Golden code.

### V. PERFORMANCE ANALYSIS

In this section, we analyze the performance of different space-time codes for a 2 × 2 and 3 × 3 MIMO-SDM systems. We compare the performance in term of bit-error rate (BER) curves versus the signal to noise ratio. At the receiver a maximum-likelihood (ML) decoder search for the codeword that minimizes the quadratic distance with the transmitted symbol.

**TABLE I:** Average number of orthogonal and non-orthogonal neighbors of \(X_i\) for different codes.

<table>
<thead>
<tr>
<th>Code</th>
<th>(N_{1_{\text{min}}})</th>
<th>(N_{2_{\text{min}}})</th>
<th>(N_{\text{min},T})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silver Code</td>
<td>6.5</td>
<td>1.5</td>
<td>8</td>
</tr>
<tr>
<td>TAST Code</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Golden Code</td>
<td>0</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Alamouti Code</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>

**A. Two-Mode Fiber Optical Channel**

In order to have a clear sight on the impact of the orthogonality of the codeword difference matrices on the error probability, we simulated four different space-time codes which are the Silver, Golden, Alamouti and TAST code. We use 4-QAM symbols to construct codewords of the Silver, Golden and TAST code (full rate codes with 2 symbols/cu), and 16-QAM symbols to construct codewords for the Alamouti code (half rate code with 1 symbol/cu).

In Fig.1 we report the performance of the Alamouti code. We notice that this code gives the same performance for different MDL values. This results is confirmed by the upper bound of equation (13).

**Fig. 1:** Bit Error Rate as a function of SNR for the Alamouti code for MDL = 6, 10 dB.

In Fig. 2 we compare the performance of the Silver, TAST and Golden code for MDL = 10 dB. we notice that the silver code outperforms the two other codes, this can be explained by the higher number of orthogonal codeword differences that the silver code has . The Golden code is the less optimal because for this code \(X_i\) has no orthogonal neighbors at \(d_{\text{min}}\). The performance of the TAST code comes between the Silver and the Golden with an average number of orthogonal neighbors equal to 4.

These results are completely different from wireless Rayleigh channel where the Golden code is the best code.

**B. Three-Mode Fiber Optical Channel**

In this subsection, we study the impact of MDL on the error probability. We use a 3 × 3 TAST code which is a full rate...
code. In Fig. 3 we report the performance of this code for MDL = 6, 10 dB. We notice that for MDL = 6 dB the TAST code absorbs all MDL. For MDL = 10 dB and at BER = 10^{-4}, the SNR penalty of the SDM optical channel to the Gaussian channel is 1.2 dB. This behavior can also be explained by the equation (31). In fact, an increasing MDL leads to increase the second term of the upper bound and hence a loss in the performance.

VI. CONCLUSION

Space time coding is a high potential technique that improves the performance of space division multiplexing systems based optical fibers. In this paper, we established an upper bound for the error probability of SDM optical channels affected by MDL. This expression allows to compare the performance of different existing space-time codes, the expression also yields a design criterion for new codes construction. According to this criterion the number of orthogonal codeword differences should be maximized. In our future work, we look forward to construct a new space-time code that satisfies this criterion and therefore absorbs all channel penalties induced by MDL.

REFERENCES