This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2654259, Journal of Lightwave Technology

Orthogonal STBC for MDL Mitigation in Mode Division Multiplexing System with MMSE Channel Estimation

Yong qiang Hei, Wen tao Li, Xiao chuan Xu, *Member, IEEE* and Ray T. Chen, *Fellow, IEEE*

***Abstract*—Mode division multiplexing (MDM) has the potential of increasing the capacity of optical fiber transmission system. However, primary impaired factor mode-dependent loss (MDL) poses fundamental performance limitations. Recently MIMO space time coding technique has provided a potential solution to mitigate the MDL and hence started receiving considerable interests. In this paper, orthogonal space time block coding (OSTBC) based MDM transmission system was investigated to test its efficiency in mitigating MDL. Considering the importance and necessity of channel estimation for coherent receivers and the paucity of literature on this subject, MMSE channel estimation process was introduced in the MDM system to test the robustness and efficiency of the OSTBC-MDM system in mitigating the MDL. Furthermore, an alternative linear decoding (LD) method was achieved by transferring the orthogonal characteristics of codes in OSTBC to the channel elements in MDM system. The LD method has the same performance as the optimal ML method, excepting that it only exhibits linear computational complexity, in terms of the number of modes and the size of constellations. Simulation results were provided to validate the efficiency and robustness of the proposed OSTBC-MDM system in mitigating the MDL in the presence of MMSE channel estimation.**

***Index Terms*—Mode division multiplexing, space time block coding, maximum likehood, channel estimation**

I. Introduction

The exponential growth of internet protocol (IP) traffic, such as cloud computing, high-quality video streaming, and mobile networking, poses significant challenges to the current communication system [1]. Nowadays, the capacity of single-mode fibers (SMF) has been rapidly approaching the fundamental Shannon limit, which cannot be surpassed by improving the current SMF systems. The potential of space division multiplexing (SDM) [2-3], using few-mode fibers (FMFs) or multi-core fibers (MCFs) has been experimentally demonstrated in achieving ultra-high spectral and spatial efficiency [1-4]. Owing to the extra degree of freedom (DOF) in the space domain, the channel capacity of an SDM system can be greatly improved. It has been shown that, by using N spatial modes of FMFs, in addition to the current multiplexing methods, the mode division multiplexing (MDM) system can achieve N-times the transmission capacity with respect to SMF [5]. By exploiting orthogonal spatial and polarization modes of FMFs to transmit independent parallel data streams [6-8], FMF based MDM system is equivalent to the multiple-input multiple-output (MIMO) system, which can greatly increase the channel capacity and spectral efficiency [9-10].

In MDM systems, the multiple spatial modes supported by the FMF have different group velocities, which causes an effect called modal dispersion (MD) that can lead to differential mode group delay (DMGD) and temporal inter-symbol interference (ISI) [11-12]. To deal with this problem, MIMO time-domain equalizers (TDEs) or frequency-domain equalizers (FDEs) can be employed [13-14]. In [15-17], common equalizer schemes were investigated and then the implementation complexity, throughput efficiency, tracking ability and the convergence time of those algorithms were discussed. While the currently used coherent transceivers equipped with linear equalizers can sufficiently address the dispersive effects, they may show limited ability in dealing with the non-unitary crosstalk effects.

On the other hand, the multiple spatial modes, supported by the FMF in MDM systems, always experience different mode-dependent gain or loss, which results in a loss of the orthogonality of the propagating modes and signal-to-noise ratio (SNR) disparities [18-19]. The modal loss disparities induce non-unitary effect, also known as mode dependent loss (MDL), which can severely significantly deteriorate the overall system capacity and BER performance [20]. Unlike MD, which does not fundamentally degrade the system performance but affects the receiver complexity, MDL poses fundamental channel capacity-limited performance, because it can reduce the number of modes available for multiplexing [21-24]. Hence, efficient measures should be taken to address the MDL issue.

Until now, plenty of research has been carried out to study the effects of MDL on MDM system performance [21-27]. In [21], it has been shown that frequency diversity in MDM system can help in mitigating the frequency dependence of the capacity. The MDM system model in [22] provides an elegant statistical representation of MDL by equating it to the eigenvalue distribution of a zero-trace Gaussian unitary ensemble. In [25], the MIMO capacity in SDM system has been systematically investigated with emphasis on system outage. The impact of MDL and the distributed noise loading on system capacity and outage has been thoroughly discussed.

This work was supported in part by the National Natural Science Foundation of China (61201135, 61201134); The 111 Project (B08038); the Fundamental Research Funds for the Central Universities (7214569601), the Shaanxi Natural Science Foundation (No. 2015JQ6245) and in part by the Research Fund of State Key Laboratory.

Y. Q. Hei is with the State Key Laboratory of ISN, Xidian University, Xi'an 710071, China (e-mail: yqhei@mail.xidian.edu.cn)

W. T. Li is with the Science and Technology on Antenna and Microwave Laboratory, Department of Electronic Engineering, Xidian University, Xi'an 710071, China (e-mail: wtli@mail.xidian.edu.cn).

X. C. Xu and R. T. Chen are with the Microelectronics Research Center, Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78758 USA (e-mail: xiaochuan.xu@utexas.edu; raychen@ uts.cc.utexas.edu).

In [26], an efficient approach for assessing the effects of MDL in SDM system, by means of MDL vector and noise degree-of-coherency vector has been proposed. In [27], a new metric, the capacity loss per mode has been proposed to quantify the impact of MDL on SDM system performance. Its advantage over the frequently used metric, the capacity ratio, has been verified. The capacity loss caused by MDL with different amplification schemes has been discussed and the error caused in this metric by different noise loadings quantified.

Recently, it has been found that MIMO detection based digital signal processing (DSP) technique offers an interesting solution to mitigate the polarization-dependent loss (PDL) in the polarization multiplexed (PolMux) system [28] and the MDL in the MDM system [29-30]. In [29], the maximum-likelihood (ML) detection and the ZF/MMSE equalizer, commonly used for non-unitary MIMO channels, have been evaluated in an MDL-impaired MDM system. It shows that for strongly coupled modes, the optimal ML detection yields a significant advantage over the ZF/MMSE equalizer. However, the ML detection suffers considerably from the exponential complexity caused by the exhaustive search. To address this issue, a sphere decoder (SD) based reduced complexity ML detection technique has been proposed in [30]. Since SD searches the constellations located only in the sphere, its complexity is greatly related to the radius of the sphere. In the worst case, it still has an exponential complexity.

Apart from the detecting techniques implemented at the receiver, advanced MIMO techniques, such as the space-time coding (STC), also can be implemented at the transmitter to better equalize the non-unitary effects of MDL and thus reduce the performance penalties. STCs, originally developed for MIMO systems to enhance transmission reliability by the code diversity gain [31-32], have demonstrated their unique advantages in mitigating PDL in PolMux systems [28, 33-36] and MDL in MDM systems [37-39]. In [34], the efficiency of Golden code (GC) and Silver code (SC), in mitigating PDL impairments, has been demonstrated. In [35-36], Meron *et al*. pointed out both GC and SC can perform well in dealing with the average 0-6dB PDL values, whereas Alamouti code(AC) can perform better in dealing with large PDL values than with low PDL values. In [37-38], linear threaded algebraic STCs (TA-STCs) have been investigated in 3- and 6-mode multiplexed MDM transmission systems, with optimal ML detection to mitigate MDL. It shows that as high as 10 dB MDL level, can be mitigated in the 6-mode MDM system. In [39], a new STC architecture with four dimensional modulation formats, has been used in a 3-mode FMF system to enhance the performance, and OSNR gains achieved.

GC, SC, TA-STC and ST-trellis codes, mentioned above, face the challenge of the optimal ML decoding suffering considerably from the exponential complexity, which is prohibitive for DSP implementation, particularly for larger number of modes or higher order modulations. Hence, to address the MDL in MDM system, it is desirable to employ STCs with sufficient diversity gain but with lower decoding complexity. On the other hand, for coherent receivers, it is important to get the channel state information (CSI) both for equalization and detection at the receiver [40]. Once the channel is known by channel estimation, the equalizers can more effectively track MIMO channels, which are more tolerant to MDL, computationally less complex and faster in convergence [11-14]. In addition, with the estimated channel matrix, it is helpful not only to compensate for the dispersion, but also to determine the MDL level. However, to the best of our knowledge, most of the extant published works on MDM MIMO systems assume that CSI is known at the receiver by channel estimation [14, 25, 29, 37-38], but they do not discuss much about the detailed process or its related effects. According to [41-42], CSI cannot be perfectly obtained by channel estimation with limited pilot symbols. In [14], least square (LS) channel estimation method was adopted in the data-aided equalization of PolMux system. In [43], intra-symbol frequency domain averaging based channel estimation method was proposed for coherent OFDM system and its effects of noise, MD, PDL and nonlinear impairments were discussed. In view of its important role in MDM system, channel estimation deserves further investigations.

In this paper, motivated by the advantage offered by Alamouti (2×2 OSTBC) in mitigating PDL [35-36], a 4×4 orthogonal space time block coding (OSTBC)-MDM system was proposed to study its efficiency in migrating the MDL. Different from the previous published works [20, 22-27] assume that perfect CSI is available when studying the effects of MDL on the capacity or BER performance of MDM systems, for this investigation, both the capacity performance and the BER performance impaired by the accumulated MDL in the presence of MMSE channel estimation were studied. Furthermore, our investigation differs from previous research on channel estimation [14,43] in that it introduces MMSE channel estimation into the MDM system to examine the robustness of the OSTBC-MDM system in mitigating the MDL. Finally, a linear decoding (LD) algorithm has been achieved by transferring the orthogonal characteristics of the OSTBC elements to the channel elements of MDM system, which could achieve the same performance as the optimal ML method in [31], but greatly reduces the complexity. More importantly, compared with the linear decoding in [31], the proposed LD totally avoids the calculation of the last term of Eqs. (7) and (8). Therefore, it greatly simplifies the signal processing burden and provides an interesting DSP solution. Simulation results show that the OSTBC-MDM system has the potential of mitigating the MDL in the presence of MMSE channel estimation.

The rest of the paper is organized as follows. Section II defines the MDM system model. Section III describes the channel estimation process in MDM system. Section IV illustrates the OSTBC-MDM transmission system. In Section V, the simulation results and discussions are provided. Section VI concludes the paper.

II. MDM Optical Transmission System

Various scenarios of MDL-impaired MDM transmission systems have been suggested in the previous works [18], [22], [29], and [44]. Since the MDM system model in [22] provides an elegant statistical representation of MDL through equivalence to the eigenvalue distribution of a zero-trace Gaussian unitary ensemble, it has been employed in this manuscript. As in previous works [18], [44], here also only a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2654259, Journal of Lightwave Technology

single polarization per spatial mode has been considered to focus solely on the modal loss disparities. The non-linear effects have been neglected and the frequency dependence of MDL has been considered. For simplicity, it has been assumed that the ISI caused by the DMD is suppressed and that the dispersion effect is compensated at the receiver DSP.

2.1. Multi-Section MDM System Model

For this investigation, the MDM transmission system model proposed by Ho and Kahn in [22] has been employed, which describes the impact of MDL accumulation in the regime of strong mode coupling. The optical link can be modeled as a concatenation of *K* short sections and in each section the propagation is a random matrix. Assuming that N orthogonal propagating modes are supported by the FMF, at angular frequency w, the overall transfer matrix is modeled by:

$$\begin{aligned}\mathbf{H}_{N\times N}^{(t)} &= \prod_{k=1}^{K}\mathbf{M}_k\left(w\right)=\prod_{k=1}^{K}\mathbf{U}_k\mathbf{\Lambda}_k\left(w\right)\mathbf{V}_k^H \\ &= \prod_{k=1}^{K}\mathbf{U}_k diag\left[\lambda_1 \quad \cdots \quad \lambda_N\right]\mathbf{V}_k^H\end{aligned} \tag{1}$$

where $\mathbf{U}_k$ and $\mathbf{V}_k$ are frequency-independent random unitary matrices, representing the mode couplings at the input and output, respectively. $\mathbf{\Lambda}_k\left(w\right)$ is a diagonal matrix representing modal propagation matrix in the *k*th section. Including both MDL and MD, its diagonal elements can be expressed by $\lambda_{i=1\cdots N}^{(k)}=\exp\left(g_i^{(k)}/2-jw\tau_i^{(k)}\right)$, where $g_i^{(k)}$ is the uncoupled modal gain in the *k*th section, measured in dB or log power gain unit, satisfying $g_1^{(k)}+\cdots+g_N^{(k)}=0$ with the root-mean-square (rms) value σ_g. If $K>>1$, the rms accumulated MDL is defined as $\xi=\sqrt{K}\sigma_g$ in the strong coupling regime. $\tau_i^{(k)}$ is the uncoupled modal group delay satisfying $\tau_1^{(k)}+\cdots+\tau_N^{(k)}=0$. In each section, the uncoupled MDL $\mathbf{g}^{(k)}$ and group delay $\boldsymbol{\tau}^{(k)}$ are modeled as frequency independent within the frequency band of interest. Multiplication of *K* matrices in (1) makes the overall MDL and group delay become frequency-dependent. By using singular value decomposition, we have $\mathbf{H}^{(t)}=\mathbf{U}_{(t)}\mathbf{\Lambda}_{(t)}\mathbf{V}_{(t)}^H$, where $\mathbf{U}_{(t)}$, $\mathbf{\Lambda}_{(t)}$ and $\mathbf{V}_{(t)}$ are all frequency-dependent. According to [22], the standard deviation （STD） of over modal gains σ_{MDL} in dB is determined by:

$$\sigma_{MDL}=\xi\sqrt{1+\xi^2/12} \tag{2}$$

With the overall channel transfer matrix established above, the $N\times 1$ received vector is obtained as follows：

$$\mathbf{y}=\mathbf{Hx}+\mathbf{n} \tag{3}$$

where $\mathbf{H}$ is the $N\times N$ channel matrix and $\mathbf{x}$ is the $N\times 1$ transmitted data signals with $\mathbf{Q}=E\left\{\mathbf{xx}^H\right\}=P_d/N\cdot\mathbf{I}_N$. P_d is the total transmit power and $\mathbf{n}$ is the independent and identically distributed (i.i.d.) additive white Gaussian circular noise (AWGN) with $\mathbf{n}\sim CN(\mathbf{0},\sigma_n^2\mathbf{I}_N)$. For convenience, the average receiver SNR is defined as $\rho_d=P_d/\sigma_n^2$. Note that in MDM system, the amplified spontaneous emission (ASE) noise generated at each amplification stage accounts for the dominant noise and is subject to MDL at each segment. Hence, the accumulated noise will become spatially colored and its coherency matrix will not be proportional to the identity. However, in [22], the authors pointed out that in the regime of strong mode coupling, the noise will become spatially white when the number of noise sources is large. Hence, for this study, the hypothesis of spatially white noise lumped at the receiver will be maintained.

2.2. Transmitter power allocation

Generally, at the transmitter the total signal power can be allocated to different modes according to the CSI before data transmission. Two cases, one with perfect CSI and the other with no CSI, are considered below:

- **Perfect CSI:** if perfect CSI is available at the transmitter, power is adaptively allocated to each mode through water-filling strategy to achieve the maximum capacity below:

$$C=E\left\{\sum_{i=1}^{N_{\min}}\log_2\left(1+P_i\,\lambda_i^2/N_0\right)\right\} \tag{4}$$

where λ_i represents the *i*-th eigenvalue of $\mathbf{H}$ and $P_i=\left[\mu-N_0/\lambda_i^2\right]^+$ is the waterfilling power allocated to the *i*-th mode. $\left[x\right]^+=\max(x,0)$ guarantees that the power value allocated to each mode cannot be less than zero. Note that the water-filling level μ should satisfy the total power constraint, i.e. $\sum_i\left[\mu-N_0/\lambda_i^2\right]^+=P_d$.

- **No CSI**: if CSI is not available at the transmitter, the power will be equally allocated to each mode. The achieved channel capacity is defined by:

$$C=\log_2\det\left(\mathrm{I}+\frac{\rho_d}{N}\mathbf{HH}^H\right) \tag{5}$$

2.3. Receiver detection

For multiplexing MDM system, at the receiver the transmitted data can be recovered by the following optimal ML decoder, as described below:

$$\mathbf{x}_{ML}=\underset{\mathbf{x}\in\Theta^N}{argmin}\left\|\mathbf{y}-\mathbf{Hx}\right\|^2 \tag{6}$$

Obviously, ML has an exponential complexity (Θ^N) with respect to the constellation size Θ and the number of modes *N* because it requires an exhaustive search over all kinds of constellation combinations.

III. Channel Estimation Process In MDM System

For MDM system with coherent detection, it is important to get the CSI for equalization and detection at the receiver [40], [45]. This can be achieved by channel estimation using the training sequences [41-42]. In this section, MMSE channel estimation is introduced in MDM system. To estimate the channel matrix $\mathbf{H}$ at the receiver, an $N\times L$ ($L\geq N$) dimensional orthogonal training sequence matrix $\mathbf{S}=\left[\mathbf{s}_1,\cdots,\mathbf{s}_L\right]$ is sent from the transmitter, satisfying $\mathbf{SS}^H=P_{tr}/N\cdot\mathbf{I}_N$. Then the received signals can be written by:

$$\mathbf{Y}_{tr}=\mathbf{HS}+\mathbf{N}_{tr} \tag{7}$$

where $\mathbf{Y}_{tr}$ represents the $N \times L$ matrix of the received signals and $\mathbf{N}_{tr}$ is the $N \times L$ matrix of the i.i.d. AWGN noise with $\mathbf{N}_{tr} \sim CN(\mathbf{0}, \sigma_n^2 \mathbf{I}_N)$. For convenience, the average training SNR is defined as: $\rho_t = P_{tr} / \sigma_n^2$. According to the MMSE channel estimation criteria [41-42], by using the received signal matrix $\mathbf{Y}_{tr}$ and the training sequence $\mathbf{S}$, the estimated $\hat{\mathbf{H}}$ can be obtained as follows:

$$\hat{\mathbf{H}} = \mathbf{Y}_{tr} \left(\mathbf{S}^H \mathbf{R}_H \mathbf{S} + \sigma_n^2 \mathbf{I} \right)^{-1} \mathbf{S}^H \mathbf{R}_H \tag{8}$$

where $\mathbf{R}_H = E\{\mathbf{H}^H \mathbf{H}\} = diag[\lambda_1, \cdots, \lambda_N]$. The estimation error matrix $\tilde{\mathbf{H}} = \mathbf{H} - \hat{\mathbf{H}}$ has zero mean and its covariance matrix is written by:

$$\mathbf{R}_{\tilde{\mathbf{H}}} = E\{\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\} = \left(\mathbf{R}_H^{-1} + P_{tr} / \sigma_n^2 N \cdot \mathbf{I}_N \right)^{-1} \tag{9}$$

Due to the orthogonal property of the MMSE estimation, $\tilde{\mathbf{H}}$ and $\hat{\mathbf{H}}$ are uncorrelated.

With the estimated channel matrix $\hat{\mathbf{H}}$, the received signal can be written by:

$$\mathbf{Y} = \mathbf{H}\mathbf{x} + \mathbf{N} = \hat{\mathbf{H}}\mathbf{x} + \tilde{\mathbf{H}}\mathbf{x} + \mathbf{N} \tag{10}$$

It can be seen that only the first term in (10) is considered as the desired signal from the transmitter and the remaining terms($\tilde{\mathbf{H}}\mathbf{x}$ and $\mathbf{n}$) are considered as interference or noise. For the uncoded MDM system, with the estimated channel matrix $\hat{\mathbf{H}}$, the ML detection at the receiver becomes：

$$\mathbf{x}_{ML} = \underset{\mathbf{x} \in \Theta^N}{argmin} \left\| \mathbf{Y} - \hat{\mathbf{H}}\mathbf{x} \right\|^2 \tag{11}$$

Then the achievable channel capacity in the presence of MMSE channel estimation can be expressed by:

$$C = E\left[\log_2 \det\left(\mathbf{I}_N + \mathbf{R}_{\tilde{\mathbf{H}}\mathbf{x}+\mathbf{n}}^{-1} \hat{\mathbf{H}}\mathbf{Q}\hat{\mathbf{H}}^H \right) \right] \tag{12}$$

where $\mathbf{R}_{\tilde{\mathbf{H}}\mathbf{x}+\mathbf{n}} = \sigma_n^2 \mathbf{I} + E\left[\tilde{\mathbf{H}}\mathbf{Q}\tilde{\mathbf{H}}^H \right]$ is the covariance matrix of the equivalent noise $\tilde{\mathbf{H}}\mathbf{x} + \mathbf{n}$. Since $E\left[\tilde{\mathbf{H}}\mathbf{Q}\tilde{\mathbf{H}}^H\right] = Tr(\mathbf{Q}\mathbf{R}_{\tilde{\mathbf{H}}}) \cdot \mathbf{I}$, we have $\mathbf{R}_{\tilde{\mathbf{H}}\mathbf{x}+\mathbf{n}} = \left(\sigma_n^2 + Tr(\mathbf{Q}\mathbf{R}_{\tilde{\mathbf{H}}}) \right) \cdot \mathbf{I}$. Then the channel capacity, in the presence of MMSE channel estimation can be further expressed by:

$$C = E\left[\log_2 \det\left(\mathbf{I}_N + \frac{P_d}{N\left(\sigma_n^2 + Tr(\mathbf{Q}\mathbf{R}_{\tilde{\mathbf{H}}}) \right)} \hat{\mathbf{H}}\hat{\mathbf{H}}^H \right) \right] \tag{13}$$

Note that in the MDM transmission system, obtaining the second order statistics $\mathbf{R}_{\tilde{\mathbf{H}}}$ of the channel estimation error matrix $\tilde{\mathbf{H}}$ is much easier than obtaining the instantaneous channel transfer matrix $\mathbf{H}$.

IV. Linear ML Detecting For OSTBC-MDM System

In the OSTBC-MDM system, the data symbols are first coded by an OSTBC before transmission, unlike in the uncoded multiplexed MDM system. Without loss of generality, it is assumed that the OSTBC matrix over a block of T_c time slots is given by $\mathbf{X}_{N \times T_c} = \left[\mathbf{A}_1 \mathbf{s} + \mathbf{B}_1 \mathbf{s}^*, \cdots, \mathbf{A}_{T_c} \mathbf{s} + \mathbf{B}_{T_c} \mathbf{s}^* \right]$, where $\mathbf{A}_t, \mathbf{B}_t, t = 1 \cdots T_c$ are the real and the image coefficient matrices, respectively. In the absence of phase noise and frequency offsets, the $N \times T_c$ received signal $\mathbf{Y}$ can be expressed by:

$$\begin{aligned} \mathbf{Y} &= \mathbf{H}\mathbf{X} + \mathbf{N} \\ &= \mathbf{H}[\mathbf{A}_1 \mathbf{s} + \mathbf{B}_1 \mathbf{s}^*, \cdots, \mathbf{A}_{T_c} \mathbf{s} + \mathbf{B}_{T_c} \mathbf{s}^*] + \mathbf{N} \\ &= \left\{ \left[\mathbf{H}\mathbf{A}_1, \cdots, \mathbf{H}\mathbf{A}_{T_c} \right] \mathbf{s} + \left[\mathbf{H}\mathbf{B}_1 \cdots, \mathbf{H}\mathbf{B}_{T_c} \right] \mathbf{s}^* \right\} + \mathbf{N} \end{aligned} \tag{14}$$

where $\mathbf{Y} = [\mathbf{y}_1 \ \cdots \ \mathbf{y}_{Tc}]$, $\mathbf{H} = [\mathbf{h}_1 \ \cdots \ \mathbf{h}_{Tc}]$ and $\mathbf{N} = [\mathbf{n}_1 \ \cdots \ \mathbf{n}_{Tc}]$. According to the OSTBC construction criterion in [31-32], $\mathbf{s}$ and $\mathbf{s}^*$ do not appear simultaneously in the same time slot. In other words, either $\mathbf{A}_t$ or $\mathbf{B}_t$ must be a zero coefficient matrix while the other is not zero, i.e., when $\mathbf{A}_t = \mathbf{0}$, $\mathbf{B}_t \neq \mathbf{0}$ and vice versa. Then, applying the following linear equivalent transformation on (14), yields:

$$\mathbf{H}_e = \begin{bmatrix} \mathbf{H}\mathbf{A}_1 + \mathbf{H}^*\mathbf{B}_1^* \\ \vdots \\ \mathbf{H}\mathbf{A}_{T_c} + \mathbf{H}^*\mathbf{B}_{T_c}^* \end{bmatrix}, \begin{cases} \mathbf{r}_t = \mathbf{y}_t, \mathbf{z}_t = \mathbf{n}_t & if \ \mathbf{A}_t \neq \mathbf{0} \\ \mathbf{r}_t = \mathbf{y}_t^*, \mathbf{z}_t = \mathbf{n}_t^* & if \ \mathbf{B}_t \neq \mathbf{0} \end{cases} \tag{15}$$

We have：

$$\mathbf{R} = \mathbf{H}_e \mathbf{s} + \mathbf{Z} \tag{16}$$

where $\mathbf{R} = [\mathbf{r}_1^H \ \cdots \ \mathbf{r}_{T_c}^H]^H, \mathbf{Z} = [\mathbf{z}_1^H \ \cdots \ \mathbf{z}_{T_c}^H]^H$. The new equivalent channel $\mathbf{H}_e$ satisfies:

$$\begin{aligned} \mathbf{H}_e^H \mathbf{H}_e &= \sum_{t=1}^{T_c} (\mathbf{H}\mathbf{A}_t + \mathbf{H}^*\mathbf{B}_t^*)^H (\mathbf{H}\mathbf{A}_t + \mathbf{H}^*\mathbf{B}_t^*) \\ &= (\sum_{t=1}^{T_c/2} \mathbf{A}_t^H \mathbf{H}^H \mathbf{H}\mathbf{A}_t + \sum_{t=T_c/2+1}^{T_c} \mathbf{B}_t^T \mathbf{H}^T \mathbf{H}^* \mathbf{B}_t^*) \\ &= \Re(\sum_{t=1}^{T_c/2} 2\mathbf{A}_t^H \mathbf{H}^H \mathbf{H}\mathbf{A}_t) = 2N \cdot Tr(\mathbf{H}^H \mathbf{H}) \cdot \mathbf{I}_N \end{aligned} \tag{17}$$

By using the linear equivalent transformation in (15), the orthogonal property of OSTBC is transferred to the new equivalent channel matrix. Furthermore, by separating the real and imaginary parts, (16) can be rewritten as:

$$\mathbf{R}_e = \mathbf{\Omega}_e \mathbf{s}_e + \mathbf{Z}_e, \mathbf{\Omega}_e = \begin{bmatrix} \Re(\mathbf{H}_e) & -\Im(\mathbf{H}_e) \\ \Im(\mathbf{H}_e) & \Re(\mathbf{H}_e) \end{bmatrix} \tag{18}$$

where $\mathbf{R}_e = [\Re(\mathbf{r}) \ \ \Im(\mathbf{r})]^T, \mathbf{s}_e = [\Re(\mathbf{s}) \ \ \Im(\mathbf{s})]^T$ and $\mathbf{Z}_e = [\Re(\mathbf{z}) \ \ \Im(\mathbf{z})]^T$. $\Re(\mathbf{x})$ and $\Im(\mathbf{x})$ represent the real and imaginary parts of $\mathbf{x}$, respectively. Then, the ML detector is further simplified as:

$$\begin{aligned} s_{ML} &= \underset{\mathbf{s} \in \Theta^{2N}}{argmin} \left\| \mathbf{R}_e - \mathbf{\Omega}_e \mathbf{s} \right\|_F^2 \\ &= \underset{\mathbf{s} \in \Theta^{2N}}{argmin} \left\| \mathbf{\Omega}_e^T \mathbf{R}_e - \mathbf{\Omega}_e^T \mathbf{\Omega}_e \mathbf{s} \right\|_F^2 \\ &= \underset{\mathbf{s} \in \Theta^{2N}}{argmin} \sum_{i=1}^{2N} \left(r_{e_i} - c_i s_i \right)^2 \\ &= \sum_{i=1}^{2N} \underset{s_i \in \Theta}{argmin} \left(r_{e_i} - c_i s_i \right)^2 \end{aligned} \tag{19}$$

Due to the convex characteristic of the quadratic function, the last equation of (19) holds. The first equation of (19) indicates that ML requires an ergodic exhaustive searching over all combinations of $\mathbf{s} \in \Theta^{2N}$. However, if we define the last equation of (19) as the linear detection (LD) method, it can achieve the same optimal solution as the ML method in [31], without sacrificing any performance. Since only parallel searching over $s_i \in \Theta$ for $2N$ real-valued individual branch is required, the overall decoding complexity is substantially reduced. Eq. (19) clearly shows how the exponential complexity of ML changes to linear complexity of the LD method. More importantly, in comparison with the linear

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2654259, Journal of Lightwave Technology

detecting method in [31], the LD method can totally avoid calculation of the last term of Eq. (7) and (8) in [31]. This can further reduce the complexity and provide an interesting DSP solution for practical implementation.

The aforementioned LD method is derived under the assumption that CSI is available at the receiver. Next, the LD method is extended to MDM system with MMSE channel estimation. With the estimated channel matrix $\hat{\mathbf{H}}$, the received signal $\mathbf{Y}$ can be expressed by :

$$\mathbf{Y} = \mathbf{HX} + \mathbf{N} = \hat{\mathbf{H}}\mathbf{X} + \tilde{\mathbf{H}}\mathbf{X} + \mathbf{N} \quad (20)$$

Similarly, the linear equivalent transformation is applied at the receiver, which gives:

$$\hat{\mathbf{H}}_e = \begin{bmatrix} \hat{\mathbf{H}}\mathbf{A}_1 + \hat{\mathbf{H}}^*\mathbf{B}_1^* \\ \vdots \\ \hat{\mathbf{H}}\mathbf{A}_{T_c} + \hat{\mathbf{H}}^*\mathbf{B}_{T_c}^* \end{bmatrix}, \tilde{\mathbf{H}}_e = \begin{bmatrix} \tilde{\mathbf{H}}\mathbf{A}_1 + \tilde{\mathbf{H}}^*\mathbf{B}_1^* \\ \vdots \\ \tilde{\mathbf{H}}\mathbf{A}_{T_c} + \tilde{\mathbf{H}}^*\mathbf{B}_{T_c}^* \end{bmatrix} \quad (21)$$

According to (17), we have:

$$\hat{\mathbf{H}}_e^H \mathbf{H}_e = \hat{\mathbf{H}}_e^H \hat{\mathbf{H}}_e + \hat{\mathbf{H}}_e^H \tilde{\mathbf{H}}_e = 2N \cdot Tr(\hat{\mathbf{H}}^H \hat{\mathbf{H}}) \cdot \mathbf{I}_N + \mathbf{E} \quad (22)$$

where $\mathbf{E} = \hat{\mathbf{H}}_e^H \tilde{\mathbf{H}}_e$ can be considered as the interference caused by the channel estimation error during the MMSE channel estimation. Multiplying the newly equivalent received signal $\mathbf{R}$ by $\hat{\mathbf{H}}_e^H$ produces:

$$\hat{\mathbf{R}} = \hat{\mathbf{H}}_e^H \hat{\mathbf{H}}_e \mathbf{s} + \hat{\mathbf{H}}_e^H \tilde{\mathbf{H}}\mathbf{s} + \hat{\mathbf{Z}} = 2N \cdot Tr(\hat{\mathbf{H}}^H \hat{\mathbf{H}})\mathbf{s} + \mathbf{E}\mathbf{s} + \hat{\mathbf{Z}} \quad (23)$$

By separating the real and the imaginary parts of the complex received symbols, we get:

$$\hat{\mathbf{R}}_e = \hat{\mathbf{\Omega}}_e \mathbf{s}_e + \tilde{\mathbf{\Omega}}_e \mathbf{s}_e + \hat{\mathbf{Z}}_e \quad (24)$$

where $\hat{\mathbf{\Omega}}_e$ and $\tilde{\mathbf{\Omega}}_e$ are similar to $\mathbf{\Omega}_e$, and $\hat{\mathbf{Z}}_e$ is similar to $\mathbf{Z}_e$. Then, the ML detector under imperfect CSI is written by:

$$\begin{aligned} s_{ML,MMSE} &= \underset{\mathbf{s}\in\Theta^{2N}}{argmin} \left\| \hat{\mathbf{R}}_e - \hat{\mathbf{\Omega}}_e \mathbf{s} \right\|_F^2 \\ &= \underset{\mathbf{s}\in\Theta^{2N}}{argmin} \left\| \hat{\mathbf{\Omega}}_e^T \mathbf{R}_e - \hat{\mathbf{\Omega}}_e^T \hat{\mathbf{\Omega}}_e \mathbf{s} \right\|_F^2 \\ &= \underset{\mathbf{s}\in\Theta^{2N}}{argmin} \sum_{i=1}^{2N} \left(\hat{r}_{e_i} - \hat{c}_i s_i \right)^2 \\ &= \sum_{i=1}^{2N} \underset{s_i\in\Theta}{argmin} \left(\hat{r}_{e_i} - \hat{c}_i s_i \right)^2 \end{aligned} \quad (25)$$

Similarly, the last equation of (25) can be defined as the LD method with MMSE channel estimation. The only difference is that the channel estimation error can result in degradation of both the capacity and the bit error rate (BER) performance.

V. Simulation Results

In this section, simulations were carried out to verify the efficiency of the OSTBC-MDM system in mitigating the accumulated MDL in the presence of MMSE channel estimation. Without loss of generality, a 4-mode MDM transmission system in [22] consisting of a concatenation of *K*=256 short sections with strong mode coupling, was employed. The gain vector $\mathbf{g}^{(k)}$ and the group delay vector $\boldsymbol{\tau}^{(k)}$ in each section were selected according to [22]. In the case of 4×4 MDM system, at the transmitter, QPSK symbols of unit energy were coded by the OSTBC with its code word matrix [31] given by:

$$\mathbf{X} = \begin{bmatrix} s_1 & -s_2 & -s_3 & -s_4 \\ s_2 & s_1 & s_4 & -s_3 \\ s_3 & -s_4 & s_1 & s_2 \\ s_4 & s_3 & -s_2 & s_1 \end{bmatrix} \quad (26)$$

The corresponding real coefficient matrices $\mathbf{A}_t$ for the four time slots are:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (27)$$

$$\mathbf{A}_3 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \mathbf{A}_4 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (28)$$

whereas its image coefficient matrices are $\mathbf{B}_1 = \mathbf{B}_2 = \mathbf{B}_3 = \mathbf{B}_4 = \mathbf{0}_{4\times4}$. Note that it may be difficult to get the OSTBC design for MDM system with a large number of modes. One potential solution is to divide the transmitted data frame into several blocks, in each of which the code word matrix is an OSTBC. In that case, a good balance between the diversity gain and the multiplexing gain can be obtained, and the LD method can still work after separating different blocks. Due to the page limit, the details of multiblock OSTBC scheme are not discussed in this paper. At the receiver, the data symbols were recovered by the proposed LD method. For comparison, the linear ZF/MMSE equalizer and the optimal ML detection in [29] for the uncoded MDM system, were also simulated to prove the advantage of OSTBC-MDM system in mitigating MDL.

To investigate the performance of OSTBC-MDM system affected by the accumulated MDL values, three cases, i.e., $\xi = 0dB$, $\xi = 5dB$ and $\xi = 10dB$, corresponding to low, intermediate and high MDL levels were considered. Assuming that the AWGN noise has unit average power, i.e., $\sigma_n^2 = 1$. For MMSE channel estimation, the length of the training sequence was selected as $L = 100$. The average training SNR was $\rho_t = 20dB$. Both BER and channel capacity performance were obtained by averaging over 10000 Monte-Carlo simulations. Within each data frame, a minimum of 200 symbols were transmitted.

- The capacity and BER performance affected by the accumulated MDL value ξ with perfect CSI

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2654259, Journal of Lightwave Technology

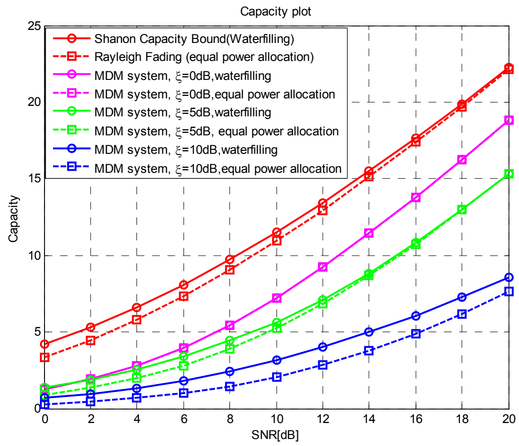

Fig. 1. Comparison of channel capacity of MDM system with different accumulated MDL values under different power allocation strategies

Figure 1 compares the channel capacity of the MDM system with different accumulated MDL values ξ under different power allocation strategies. The Shannon channel capacity bound, which can be obtained by using waterfilling power allocation strategy under the i.i.d Rayleigh fading channel, was plotted as a reference. For comparison, the capacity achieved by equal power allocation strategy without knowing the Rayleigh fading CSI was also presented. In this figure, a reduction in capacity can be seen for the MDM system with increase in ξ. When $\xi = 0dB$, the capacity achieved by waterfilling is the same as that achieved by equal power allocation. This is probably because, the disparities between different eigenvalues of the modes in the MDM system were so small that waterfilling could not fully develop its effect. With increase in the accumulated MDL values ξ, the channel capacity of the MDM system gradually decreases, and the gap between the capacity achieved by waterfilling and that achieved by equal power allocation becomes larger.

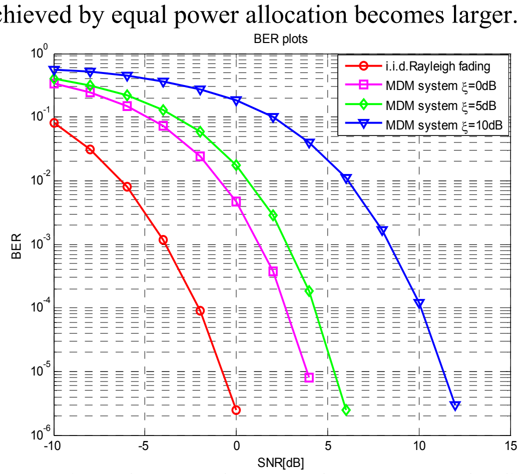

Fig. 2. Comparison of BER performance of MDM system for different accumulated MDL values with LD method

Figure 2 shows the BER performance of OSTBC-MDM system, affected by the accumulated MDL value ξ. It can be discovered that the BER performance has been significantly improved by OSTBC. Considering the large accumulated MDL value $\xi = 10dB$ case, for example, a better BER performance, i.e., BER=10^{-5} at SNR=12 dB, could be achieved, which proves the benefit of OSTBC in averaging the losses, experienced by the mode multiplexed data symbols.

- Effect of accumulated MDL ξ on the capacity and BER performance with MMSE channel estimation

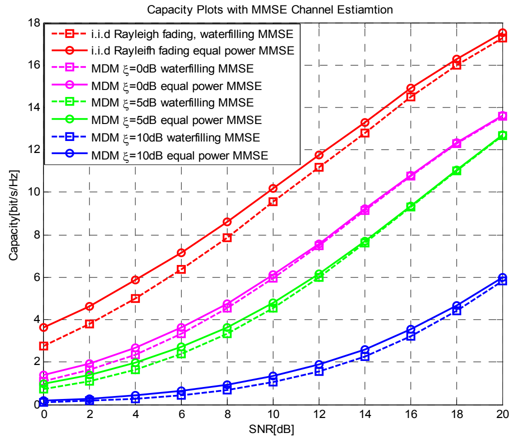

Fig. 3. Effects of accumulated MDL on the channel capacity of MDM system in the presence of MMSE channel estimation. $\rho_t = 20dB$

Figure 3 illustrates the channel capacity performance of the MDM system affected by the accumulated MDL value ξ, in the presence of MMSE channel estimation. For comparison, the capacities achieved under i.i.d. Rayleigh fading with different power allocation strategies, in the presence of MMSE channel estimation, were also provided. Compared with the perfect CSI case (see Fig.1), a reduction in channel capacity was caused due to channel estimation error. Considering that SNR=20dB, for example, when $\xi = 0dB$, $\xi = 5dB$, and $\xi = 10dB$, in the presence of MMSE channel estimation, the capacity loss is 4bit/s/Hz, 2bit/s/Hz, and 2bit/s/Hz respectively, as compared to the respective capacities achieved by waterfilling power allocation with perfect CSI. These results indicate that the channel estimation error can further degrade the channel capacity of the MDL-impaired MDM system.

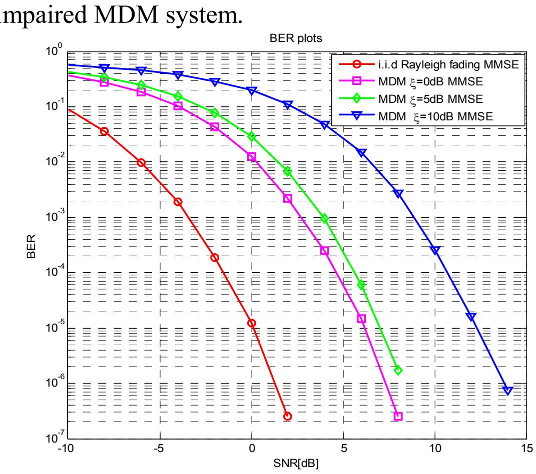

Fig. 4. Comparison of BER performance of OSTBC-MDM system with LD method for different accumulated MDL values in the presence of MMSE channel estimation. $\rho_t = 20dB$

Figure 4 shows the BER performance of MDM system, affected by MDL values in the presence of MMSE channel estimation. From this figure, it can be seen that inaccuracy of channel estimation causes further degradation of BER performance. Compared with the perfect CSI case (see Fig.2), a small SNR penalty was sacrificed to compensate for channel

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2654259, Journal of Lightwave Technology

estimation error. Nevertheless, the diversity provided by OSTBC is sufficient to combat the performance degradation caused by ξ and channel estimation error. Considering that $\xi = 0dB$, $\xi = 5dB$ and $\xi = 10dB$, for example, to achieve a target BER=10^{-5}, SNRs of 4dB, 5dB and 11dB were required respectively for the perfect CSI case, in comparison to the corresponding SNRs of 6dB, 7dB and 12dB required for the MDM system with channel estimation. These results prove the efficiency of the OSTBC in mitigating the MDL in the presence of MMSE channel estimation.

- The efficiency of OSTBC-MDM system against the MDL with MMSE channel estimation.

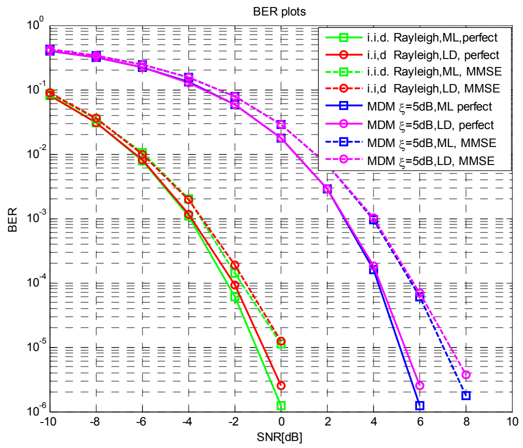

Fig. 5. Comparison of BER performance of OSTBC- MDM system between LD and ML methods under i.i.d Rayleigh fading channel in the case of perfect CSI and MMSE channel estimation

Figure 5 makes a comparison of BER performance between the linear decoding (LD) and the ML methods under different scenarios. A good agreement between the LD and the ML methods can be seen in the case of both perfect CSI scenario and MMSE channel estimation. These simulation results can validate the optimal characteristic of the LD method mentioned previously.

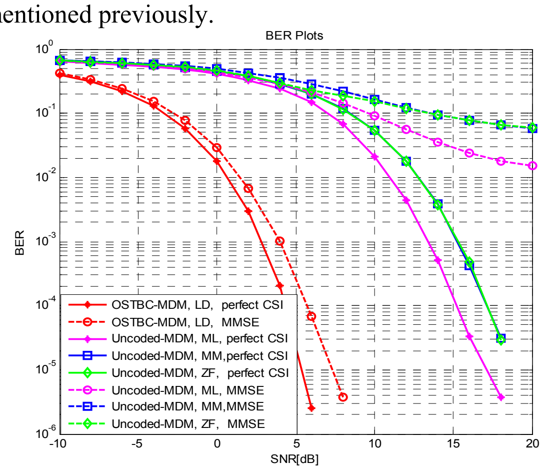

Fig. 6. Comparison of BER performance between OSTBC-MDM and uncoded-MDM systems, decoded by different detecting methods with and without MMSE channel estimation $\xi = 5dB$ $\rho_t = 20dB$

Figure 6 compares BER performance between the OSTBC-MDM and the uncoded MDM systems, decoded by different detection methods (LD, ML, MM, ZF) with and without MMSE channel estimation. The uncoded MDM system refers to the system that the data symbols at the transmitter are not coded by any STC before transmission. Note that, in the legend, MM represents the MMSE equalization method, and MMSE denotes the MMSE channel estimation. The advantage of OSTBC-MDM system over the uncoded-MDM system in migrating MDL can be clearly seen, especially in the presence of MMSE channel estimation. For OSTBC-MDM system, in the presence of MMSE channel estimation, still a better BER performance was achieved. Only a small SNR penalty was sacrificed to compensate for the degradation caused by channel estimation error. For the uncoded MDM system without STC, the ML method has the best BER performance, followed by the MM and the ZF methods. However, in the presence of MMSE channel estimation, even with the ML detecting method, the BER performance is rather limited. For example, at SNR=20dB, BER of only 10^{-2} could be achieved, which is far below the accepted BER level. These results validate the benefit of OSTBC in mitigating the MDL in the presence of channel estimation.

- **Complexity Evaluation**:

For practical implementation of MDM system, one primary factor that needs to be considered is the computational complexity of the detecting schemes at the receiver, which to a large extent determine whether those detecting schemes can be adopted or not. According to [16], the complexity of a detecting scheme can be roughly estimated by the number of complex multiplications involved. For the uncoded MDM system, a complexity of N^2M^N is required for ML detection, and $N^2 + N^3$ for the MMSE/ZF method, where N is the number of modes and M is the size of the constellations [29]. For full rate STC-MDM system, usually at the receiver, a vectorization is applied on the received signal to generate the new equivalent channel of size $NT \times NT$, where T is STC code length. Then, the ML detection requires a complexity of $(NT)^2 M^{NT}$ to recover the transmitted data. For the OSTBC-MDM system discussed in this paper, the LD method can achieve the same performance as the ML method, but requires a complexity of only $(NT)^2 M$. For convenience, comparison of the decoding complexities of different STC schemes (full rate STC, OSTBC and uncoded) for $N \times N$ MDM system is provided in Table I, in which, for simplicity, QPSK is assumed for all the modulated symbols.

Table I: Comparison of the decoding complexities of different STC schemes for $N \times N$ MDM system

STC schemes (decoding)	Full rate STC (ML)	OSTBC (LD)	Uncoded (ML)
complexity / modes	$(NT)^2 M^{NT}$	$(NT)^2 M$	$(N)^2 M^N$
2-mode	4096	64	64
3-mode	2.1×10^7	576	576
4-mode	1.1×10^{12}	1024	4096
6-mode	6.1×10^{24}	9216	1.5×10^5

From Table I, it can be seen that when full-rate STCs are employed in an $N \times N$ MDM system, the decoding complexity increases exponentially with the number of modes. It will become intractable for the MDM system with a large number of modes. Compared with the full rate STCs, OSTBC has a distinct advantage because of linear decoding at the receiver, which can significantly reduce the complexity at the sacrifice

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JLT.2017.2654259, Journal of Lightwave Technology

of a bit rate penalty. Since LD grows linearly with respect to the constellations size, it provides an interesting DSP solution for the MDM system. For uncoded multiplexing MDM system, the optimal ML decoding has nearly the same decoding complexity as the LD method for OSTBC-MDM system, but it shows a limited ability in mitigating the MDL.

VI. Conclusion

In this paper, we have proposed orthogonal space time block coding (OSTBC) assisted MDM system to migrate the MDL with MMSE channel estimation. Effects of the accumulated MDL on the capacity and BER performance of the MDM system in the presence of MMSE channel estimation have been analyzed. Furthermore, due to the benefit of OSTBC, a linear decoding method has been achieved at the receiver for the MDM system, which can achieve the same performance as the optimal ML method, but significantly reduces the decoding complexity. Since only growing linearly with respect to the size of constellations, it provides an interesting DSP solution for MDM system. Simulation results have demonstrated the unique advantage of OSTBC in mitigating the MDL of MDM system in the presence of channel estimation error.

In this paper, we have employed the lumped noise loading model in which the spatially isotropic noise is loaded at the receiver, as was done in previous works [18], [20], [22], [30], and [38]. However, this model is a little different from the distributed noise loading model in [25-27], in which the noise is assumed to be loaded at each inline amplifier so that the effects of MDL on the distributed ASE noise can be quantified. It is important to point out that these two different noise loading models have slightly different capacity performance. Since the distributed ASE noise is subject to the MDL at each segment, the accumulated amplification noise is no longer spatially white [46-48]. According to our simulation results, larger capacities can be achieved by distributed noise loading model than by lumped noise loading model, which agrees well with the conclusion achieved in [27]. This is probably because the distributed ASE noise is also polarized by MDL and reduced noise enhancement is incurred by equalization at the receiver. Therefore, compared with the distributed ASE noise model, the lumped noise model would overestimate the MDL-induced penalties.

References

[1] R. G. H. V. Uden, R. A. Correa, E. Antonio-Lopez, *et al.*, “Ultra-high-density spatial division multiplexing with a few-mode multicore fiber,” *Nature Photonics*, vol. 8, pp. 865-870, Nov. 2014.

[2] P. J. Winzer: Making spatial multiplexing a reality, *Nature Photonics*, vol. 8, pp.345-348, Apr. 2014.

[3] D. J. Richardson, J. M. Fini, and L. E. Nelson, “Space division multiplexing in optical fibres,” *Nature Photonics.*, vol. 7, no. 5, pp. 354–362, May 2013.

[4] R. J. Essiambre, R. Ryf, N. K. Fontaine, *et al*., “Breakthroughs in photonics 2012: Space-division multiplexing in multimode and multicore fibers for high-capacity optical communications,” *IEEE Photonics Journal*, vol. 5, no. 2, Apr. 2013.

[5] H. Huang, G. Milione, M. P. J. Lavery, *et al*., “Mode division multiplexing using an orbital angular momentum mode sorter and MIMO-DSP over a graded-index few-mode optical fiber,” *Scientific Report*, DOI:10.1038/srep14931, Sep. 2015.

[6] N. Ahmed, Z. Zhao, L. Li. *et al*, “Mode Division Multiplexing of Multiple Bessel-Gaussian Beams Carrying Orbital-Angular-Momentum for Obstruction-Tolerant Free-Space Optical and Millimetre-Wave Communication Links,” *Scientific Report*, DOI: 10.1038/srep22082, Mar. 2016.

[7] R. Ryf, S. Randel, A. H. Gnauck, *et al*. “Mode-division multiplexing over 96 km of few-mode-fiber using coherent 6× 6 MIMO processing”. *IEEE Journal of Lightwave Technology*, vol. 30, no. 4, pp.521–531, 2012.

[8] P. J. Winzer, “Spatial Multiplexing in Fiber Optics: The 10X Scaling of Metro/Core Capacities,” *Bell Labs Technical Journal*, vol.19, pp. 22-30, Sep. 2014.

[9] P. J. Winzer, “Spatial multiplexing: the next frontier in network capacity scaling,” *in Proc. 39th European Conference and Exhibition on Optical Communication (ECOC)*, London, UK, Sep. 2013, pp. 1-4.

[10] T. Mori, T. Sakamoto, M. Wada, *et al*., “Few-Mode fibers supporting more than two LP modes for mode-division-multiplexed transmission with MIMO DSP,” *Journal of Lightwave Technology*, vol. 32, no. 14, pp. 2468-2479, July 2014

[11] S. O. Arik, J. M. Kahn, and K. P. Ho, “MIMO Signal Processing for Mode-Division Multiplexing: An overview of channel models and signal processing architectures,” *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 25-34, Mar. 2014.

[12] S. Ö Arık, D. Askarov and J. M. Kahn, “Adaptive frequency-domain equalization in mode-division multiplexing systems,” *IEEE Journal of Lightwave Technology*, vol. 32, no. 10, pp. 1841-1852, May 2014

[13] N. Bai and G. Li, “Adaptive frequency-domain equalization for mode division multiplexed transmission,” *IEEE Photonics Technology Letters*, vol. 24, no. 21, pp. 1918–1921, Nov. 2012

[14] M. Kuschnerov, M. Chouayakh, K. Piyawanno, *et al*., “Data-aided versus blind single-carrier coherent receivers, " *IEEE Photonics Journal*, vol. 2, no. 3, pp. 387-403, Jun. 2010

[15] S. O. Arik, D. Askarov, and J. M. Kahn, “Effect of mode coupling on signal processing complexity in mode-division multiplexing,” *IEEE Journal of Lightwave Technology*, vol. 31, no. 3, pp. 423–431, Feb. 2013.

[16] B. Inan, *et al*., “DSP complexity of mode-division multiplexed receivers,” *Optics Express*, vol. 20, no. 10, pp. 10859–10869, May. 2012.

[17] B. Inan, Y. Jung, V. Sleiffer, M. Kuschnerov, *et. al,* “Low computational complexity mode division multiplexed OFDM transmission over 130 km of few mode fiber,” *Optical Fiber Communication Conference*, Optical Society of America, 2013.

[18] S. Warm and K. Petermann, “Splice loss requirements in multi-mode fiber mode-division-multiplex transmission links,” *Optics Express*, vol. 21, no. 1, pp. 519-532, Jan. 2013

[19] A. Andrusier, M. Shtaif, C. Antonelli, and A. Mecozzi, “Characterization of mode-dependent loss in SDM systems,” *Proc. OFC*, Paper Th1J.2, 2014.

[20] K. Guan, P. J. Winzer, and M. Shtaif, “BER performance of MDL-impaired MIMO-SDM systems with finite constellation inputs,” *IEEE Photonics Technology Letters*, vol. 26, no. 12, pp. 1223–1226, Jun. 2014.

[21] K. P. Ho and J. M. Kahn, “Frequency diversity in mode-division multiplexing systems,” *IEEE Journal of Lightwave Technology*, vol. 29, no. 24, pp. 3719–3726, Dec. 15, 2011.

[22] K. P. Ho and J. M. Kahn, “Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing,” *Optics Express*, vol. 19, no. 17, pp. 16612–16635, Aug. 2011.

[23] K. P. Ho and J. M. Kahn, “Statistics of group delays in multimode fiber with strong mode coupling,” *IEEE Journal of Lightwave Technology*, vol. 29, no. 21, pp. 3119–3128, 2011.

[24] K. P. Ho, and J. M. Kahn. “Linear Propagation Effects in Mode-Division Multiplexing Systems,” *IEEE Journal of Lightwave Technology*, vol. 32, no. 4, pp.614-628, Feb, 2014

[25] P. J. Winzer and G. J. Foschini, “MIMO capacities and outage probabilities in spatially multiplexed optical transport systems,” *Optics Express*, vol. 19, no. 17, pp.16680–16696, 2011.

[26] A. Andrusier, M. Shtaif, C. Antonelli, and A. Mecozzi, “Assessing the Effects of Mode-Dependent Loss in Space-Division Multiplexed Systems,” *IEEE Journal of Lightwave Technology*, vol.32, no.7, pp. 1317–1322, April, 2014.

[27] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer “Modeling and performance metrics of MIMO-SDM systems with different amplification schemes in the presence of mode-dependent loss,” *Optics*

Express, vol.23, no.3, pp. 2203–2219, Jan. 2015.

[28] S. Mumtaz, G. R. B. Othman and Y. Jaouën, "PDL mitigation in PolMux OFDM systems using Golden and Silver Polarization-Time codes," *Optical Fiber Communication* (OFC), *collocated National Fiber Optic Engineers Conference(OFC/NFOEC), 2010 Conference on*, San Diego, CA, 2010, pp. 1-3.

[29] A. Lobato, F. Ferreira, B. Inan, *et al.*, "Maximum-likelihood detection in few-mode fiber transmission with mode-dependent loss," *IEEE Photonics Technology Letters*, vol. 25, no. 12, pp. 1095-1098, Jun. 2013.

[30] A. Lobato, J. Rabe, F. Ferreira, *et al.*, "Near-ML detection for MDL-impaired few-mode fiber transmission," *Optics Express*, vol. 23, no. 8, pp. 9589–9601, Apr. 2015.

[31] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-Time Block Codes from Orthogonal Designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456-1467, Mar. 1999.

[32] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.

[33] E. Awwad, Y. Jaouen, and G. R.-B. Othman, "Polarization-time coding for PDL mitigation in long-haul PolMux OFDM systems," *Optics Express*, vol. 21, no. 19, pp. 22773-22790, Sept. 2013.

[34] S. Mumtaz, G. B. Othman and Y. Jaouen, "Space-time codes for optical fiber communication with polarization multiplexing," *Communications (ICC), 2010 IEEE International Conference on*, Cape Town, South Africa, 2010, pp. 1-5.

[35] A. Andrusier, E. Meron, M. Feder, and M. Shtaif, "Optical implementation of a space-time-trellis code for enhancing the tolerance of systems to polarization-dependent loss," *Optics Letter*, vol. 38, no. 2, pp. 118-120, Jan. 2013.

[36] E. Meron, A. Andrusier, M. Feder, and M. Shtaif, "Use of space–time coding in coherent polarization-multiplexed systems suffering from polarization-dependent loss," *Optics Letters*, vol. 35, no. 21, pp. 3547-3549, Nov. 2010.

[37] E. Awwad, G.R-B. Othman, Y. Jaouen and Y. Frignac, "Space-time codes for mode-multiplexed optical fiber transmission systems," presented at the *Advanced Photonics for Communications*, San Diego, CA, USA, 2014, Paper SM2D.4.

[38] E. Awwad, G. R. B. Othman and Y. Jaouen, "Space-time coding schemes for MDL-Impaired mode-multiplexed fiber transmission systems," *IEEE Journal of Lightwave Technology*, vol. 33, no. 24, pp. 5084-5094, Dec. 2015.

[39] C. Okonkwo, R. van Uden, H. Chen, H. de Waardt, and T. Koonen, "Advanced coding techniques for few mode transmission systems," *Optics Express*, vol. 23, no. 2, pp. 1411-1420, Jan. 2015.

[40] H. Meyr, M. Moeneclaey, and S. Fechtel, "Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing." Hoboken, NJ: Wiley, 1997.

[41] M. Biguesh, and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884-893, Mar. 2006.

[42] A. Soysal, and S. Ulukus, "Joint channel estimation and resource allocation for MIMO Systems–Part I: single-user analysis," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 624-631, Feb. 2010.

[43] Xiang Liu and Fred Buchali. "Intra-symbol frequency-domain averaging based channel estimation for coherent optical OFDM", *Optics Express*, vol. 16, no. 26, pp. 21944-21957, Dec. 2008.

[44] A. Juarez, E. Krune, S. Warm, C. Bunge, and K. Petermann, "Modeling of mode coupling in multimode fibers with respect to bandwidth and loss," *IEEE Journal of Lightwave Technology*, vol. 32, no. 8, pp. 1549–1558, Apr. 2014.

[45] J. G. Proakis and M. Salehi, "Digital Communications," 5th ed. New York, NY, USA: McGraw-Hill, 2008.

[46] P. J. Winzer and G. J. Foschini, "Mode division multiplexed transmission systems," *In Optical Fiber Communication Conference (OFC),* San Francisco, CA, USA, 2014, paper Th1J.1.

[47] P. J. Winzer, H. S. Chen, R. Ryf, K. Guan, and S. Randel, "Mode-dependent loss, gain, and noise in MIMO-SDM systems", *in Proc. European Conference and Exhibition on Optical Communication (ECOC),* Cannes, France, 2014, Paper Mo.3.3.2.

[48] P. J. Winzer and G. J. Foschini, "Optical MIMO-SDM system capacities," *In Optical Fiber Communication Conference (OFC),* San Francisco, CA, USA, 2014, paper Th1J.1.

Yongqiang Hei received the B.Sc. and PHD. degrees in telecommunication engineering from the Xidian University, Xi'an, China, in 2006 and 2010, respectively. He is now an associate professor of the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China. His research interests include MIMO, cognitive radio, resource allocation in wireless communication, mode division multiplexing, MIMO antenna design.

Wentao Li received the B.Sc. and and M.S. degrees in electronic engineering, Ph.D. degrees from the Xidian University, Xi'an, China, in 2006 and 2010, respectively. She is now an associate professor of the Science and Technology on Antenna and Microwave Laboratory, Department of Electronic Engineering, Xi'an, China. Her research interests include conformal antenna array, phased array design; flexible antenna design; beamforming algorithm, evolutionary algorithm

Xiaochuan Xu (M'15) received the B.Sc. and M.S. degrees in electrical engineering from the Harbin Institute of Technology, Harbin, China, in 2006 and 2009, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Texas, Austin, TX, USA, in 2013. His research interests include flexible photonics, nonlinear optics, fiber optics, and silicon photonics.

Ray T. Chen (M'91–SM'98–F'04) received the B.S. degree in physics from the National Tsing Hua University, Hsinchu City, Taiwan, in 1980, the M.S. degree in physics in 1983, and the Ph.D. degree in electrical engineering in 1988, both from the University of California, Sacramento, CA, USA. He holds the Keys and Joan Curry/Cullen Trust Endowed Chair at The University of Texas Austin. He is the Director of the Nanophotonics and Optical Interconnects Research Lab, at the Microelectronics Research Center.

He is also the Director of the newly formed AFOSR MURI-Center for Silicon Nanomembrane involving faculty from Stanford, UIUC, Rutgers, and UT Austin. He joined UT Austin in 1992 to start the optical interconnect research program. From 1988 to 1992, he was a Research Scientist, Manager, and the Director of the Department of Electro-Optic Engineering, Physical Optics Corporation in Torrance, CA, USA. He was the CTO, Founder, and Chairman of the Board of Radiant Research, Inc. from 2000 to 2001, where he raised 18 million dollars A-Round funding to commercialize polymer-based photonic devices involving over 20 patents, which were acquired by Finisar in 2002, a publicly traded company in the Silicon Valley (NASDAQ:FNSR). He also was the Founder and Chairman of the Board of Omega Optics Inc. since its initiation in 2001. Omega Optics has received over five million dollars in research funding.

His research work has been awarded more than 120 research grants and contracts from such sponsors as Army, Navy, Air Force, DARPA, MDA, NSA, NSF, DOE, EPA, NIH, NASA, the State of Texas, and private industry. The research topics are focused on three main subjects: 1) Nano-photonic passive and active devices for sensing and interconnect applications, 2) Thin film guided-wave optical interconnection and packaging for 2-D and 3-D laser beam routing and steering, and 3) True time delay wide band phased array antenna (PAA). Experiences garnered through these programs in polymeric and semi-conducting material processing and device integration are pivotal elements for his research work.

His group at UT Austin has reported its research findings in more than 780 published papers, including over 85 invited papers. He holds 25 issued patents. He has chaired or been a program-committee member for more than 110 domestic and international conferences organized by IEEE, The International Society of Optical Engineering, OSA, and PSC. He has served as an Editor, Co-editor or Coauthor for more than 20 books. He has also served as a Consultant for various federal agencies and private companies and delivered numerous invited talks to professional societies. He is a Fellow of the OSA and SPIE. He received the 1987 UC Regent's Dissertation Fellowship and the 1999 UT Engineering Foundation Faculty Award, for his contributions in research, teaching and services. He also received the 2008 IEEE Teaching Award, and the 2010 IEEE HKN Loudest Professor Award, 2013 NASA Certified Technical Achievement Award for contribution on moon surveillance conformable PAA. During his undergraduate years at the National Tsing Hua University, he led the 1979 university debate team to the Championship of the Taiwan College-Cup Debate Contest.

He has supervised and graduated 46 Ph.D. students from his research group at UT Austin