


# Long-Haul Mode Multiplexing Transmission Enhanced by Interference Cancellation Techniques Based on Fast MIMO Affine Projection

Kohki Shibahara , Member, IEEE, Takayuki Mizuno, Member, IEEE, and Yutaka Miyamoto, Member, IEEE

(ECOC 2019 Invited Paper)

**Abstract**—In space division multiplexing (SDM) transmission utilizing mode division multiplexing (MDM) technologies, transmission performance of each spatial channel is highly dependent on a spatial mode that delivers information symbols over few-mode fibers (FMFs) or multi-mode fibers (MMFs). This still remains as a dominant barrier to realize a future deployable long-haul MDM transport systems. In this article, we review the transmission issues we face when transmitting MDM signals over long distances. We also present the review of key techniques that we have recently developed to overcome mode-relevant linear phenomena. By developing the projection technique for a signal detection application in multiple-input multiple-output (MIMO) systems, a novel linear filtering adaptation technique is proposed to enhance the performance of our interference canceling scheme with decreased complexity. Experimental evaluation using dense SDM transmission results over multicore FMFs verifies a fast convergence property brought by the proposed projection technique.

**Index Terms**—Space division multiplexing (SDM), differential mode delay (DMD), mode dependent loss (MDL), affine projection algorithm (APA).

## I. INTRODUCTION

IN THE development of space division multiplexing (SDM) technologies in the last decade, substantially increased transmission capacity has been attained by utilizing core and/or mode multiplexing techniques. This was attributed to the advantages of SDM fibres that parallelize information-carrying optical paths, and hence accommodate many spatial channels with extremely increased aggregate spectral efficiency (SE). An approach using mode division multiplexing (MDM) over multi-mode fibres (MMFs) or few-mode fibres (FMFs) is an attractive solution for enhancing SE under optical bandwidth limitations in a single fibre with a standard cladding diameter of 125  $\mu\text{m}$ .

Fig. 1 summarizes recent progress in long-haul and/or high-capacity MDM transmission experiments with the transmission

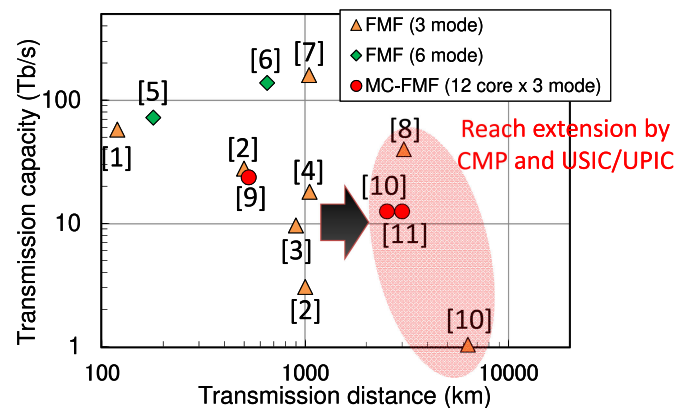


Fig. 1. Progress in MDM transmission experiments with the reach exceeding 100 km over FMFs or MC-FMFs.

reach over 100 km, including those using FMFs [1]–[8] or multicore (MC) FMFs [9]–[11]. To date, significant studies have been reported demonstrating high-capacity MDM transmission with a capacity exceeding 100 Tb/s [6], [7], [12]–[15] and long-haul MDM transmission experiments over distances more than 1000 km [6], [10], [16], [17]. Recently, by developing transmission techniques designed to achieve high-capacity long-haul MDM transmission, we have demonstrated the first ever ultra-long-haul transoceanic-class MDM transmission over FMF with reach of 6300 km [10]. The results of the longest dense SDM (DSDM) transmission over 12-core  $\times$  3-mode MC-FMF were also presented with transmission reach over 2500 km [10] and 3000 km [11] with a SE-distance product of 301398 b/s/Hz·km, which is the highest in the MC-FMF transmission ever reported. By fully utilizing the C-band, we achieved the long-haul high-capacity MDM transmission with 40.2-Tb/s capacity over 3060 km [8].

This paper starts with a discussion of MDM transmission issues and then describes key techniques that we have recently developed with the aim to transmit MDM signals over long distances. Particularly, interference cancellation techniques are briefly reviewed with a MDM-MIMO system architecture. We further extend the work presented in [18], by proposing a new computationally efficient projection technique for MDM-MIMO signal detection, with which the impact of mode

Manuscript received December 6, 2019; revised May 11, 2020 and June 3, 2020; accepted June 3, 2020. Date of publication June 10, 2020; date of current version September 15, 2020. (Corresponding author: Kohki Shibahara.)

The authors are with the NTT Network Innovation Laboratories, NTT Corporation, Kanagawa 239-0847, Japan (e-mail: kouki.shibahara.nv@hco.ntt.co.jp; takayuki.mizuno.zp@hco.ntt.co.jp; yutaka.miyamoto.fb@hco.ntt.co.jp).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JLT.2020.3001313

dependent loss (MDL) on linear filter adaptation techniques is experimentally investigated using DSDM transmission experiment results.

The rest of this paper is organized as follows: Section II discusses transmission issues faced in long-haul MDM transmission. This is followed by an overview of key transmission techniques that we have recently developed. Section III describes proposed interference cancellation techniques and their characteristics in applying MDM-MIMO systems. Section IV includes a practical linear filter design for MDM transmission together with a proposal of a reduced-complexity projection technique for linear MDM-MIMO detectors. Section V provides an experimental verification of the proposed projection technique using DSDM transmission data. Finally, Section VI concludes this paper with a summary.

*Notations:*  $(\cdot)^T$ ,  $(\cdot)^H$ ,  $(\cdot)^*$ , and  $(\cdot)^{-1}$  denote the matrix transpose, the matrix complex-conjugate transpose, complex conjugate, and the matrix inverse, respectively.  $\text{Re}(\cdot)$  denotes the real part of a matrix. The portion of the vector  $\mathbf{a}$  comprising  $i_1$ -th to  $i_2$ -th components is denoted by  $\mathbf{a}_{|i_1:i_2}$ . Likewise, the portion of the matrix  $\mathbf{A}$  comprising components from  $i_1$ -th to  $i_2$ -th rows and from  $j_1$ -th to  $j_2$ -th columns is denoted by  $\mathbf{A}_{|i_1:i_2,j_1:j_2}$ .

## II. ISSUES AND KEY TECHNIQUES IN LONG-HAUL MDM TRANSMISSION

In earlier MDM transmission experiments, achievable reach over FMFs were restricted to around 1000 km [2], [4]. This was largely due to mode dependent transmission properties that significantly deteriorate transmission performance with increased transmission reach. In particular, one dominant performance-limiting factor observed in long-haul MDM transmission is differential mode delay (DMD), which causes spreading of the impulse response width in the time domain as a consequence of multiple mode coupling along FMF transmission lines. With increased transmission reach, DMD accumulation exhibits a square-root growth in a strongly-coupled regime [19]. In our previous experimental observations where inter-mode mixing was expected to be weak in FMFs but intermediate at FM splicing/connection points, the required equalizer length (which is used as a measure of the DMD accumulation along the link) showed almost linear growth as a function of transmission distances up to 2000 km [10]. In either mode-coupling regime, an impulse response width after MDM transmission over a few thousands of km would be broadened by as much as several tens of ns (even with the use of graded-index FMFs), corresponding to weight vector length of several thousands in MIMO-DSP employing linear detection techniques. Moreover, since the DMD characteristic including wavelength dependency is highly sensitive to fibre parameters such as refractive index profiles [20], an approach of managing DMD along FMF links over wide optical bandwidth (e.g., C or L bands) would be technically challenging. Another phenomenon inherent to optical MDM transmission is MDL that mainly arises from FM components including multiplexers/demultiplexers (MUX/DEMUX), FM amplifiers, and splicing/connection points between FMF segments. MDL imposes a different loss on each spatial channel,

and decreases achievable MIMO capacity in optical MIMO link systems [21], [22]. In terms of practical use, the difficulty in dealing with MDL lies in MIMO-DSP, because a MIMO channel orthogonality is lost due to the MDL's presence, which generally makes signal detection complicated due to a high intermodal correlation in received signals. We have experimentally showed that long-haul MDM transmission yields MDL accumulation as large as 10 dB in single-core FMF transmission, and 15 dB in multi-core (MC) FMF transmission, which is not the case with polarization dependent loss in conventional transmission over single mode fibers.

To relax the MIMO-DSP requirement and achieve robust and reliable MDM transmission over long distances, we proposed a new MDM transmission scheme that induces deliberate mode interchange in each span via a pair of mode MUX/DEMUX, termed cyclic mode permutation (CMP) [10]. A schematic of a MDM-MIMO system applying a CMP scheme is illustrated in Fig. 2. In a transmission strategy with a CMP scheme, information symbols delivered on each spatial channel would periodically experience all optical paths in a SDM fiber being characterized by each spatial mode. Furthermore, a CMP scheme is expected to stimulate mode mixing process, because some specific spatial modes would effectively couple with each other when they are mutually degenerated. Indeed, our experimental evaluation results revealed that the CMP scheme enabled DMD-unmanaged long-haul MDM transmission where the broadening of the impulse response width induced by DMD was suppressed by a factor of up to 70% [10]. This CMP technique was further applied to high-capacity long-haul MDM transmission, achieving full C-band 3060-km 3-mode MDM transmission across the 4.4-THz optical bandwidth with the capacity of 40.2 Tb/s [8].

To mitigate MDL impact, we also proposed a novel MIMO-DSP technique [10], [23] where interference cancellation is successively performed in conjunction with coding techniques. This is also shown in Fig. 2. The proposed interference canceler will be described in detail in Section III. The advantage brought by this is an improvement of post-processing signal-to-interference-plus-noise ratio (SINR) under the MDL's presence. With this key technique, we achieved the longest DSDM transmission yet reported over 12-core  $\times$  3-mode MC-FMF with the transmission distances of 2500 km [10] and 3000 km [11].

## III. INTERFERENCE CANCELLATION TECHNIQUES FOR MDM-MIMO SYSTEMS

### A. System Model

We consider an MDM system where  $N_T$  data streams (including streams delivered on polarization modes) are detected by  $N_R$  receivers after propagation over FMF transmission links. Denoting  $x_i$  and  $y_i$  as the  $i$ -th transmitted symbol and the  $i$ -th received symbol, respectively, the end-to-end signal propagation in a linear transmission regime is described by an MIMO model:

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{n} \\ &= \sum_{i=1}^{N_T} \mathbf{h}_i x_i + \mathbf{n}, \end{aligned} \quad (1)$$

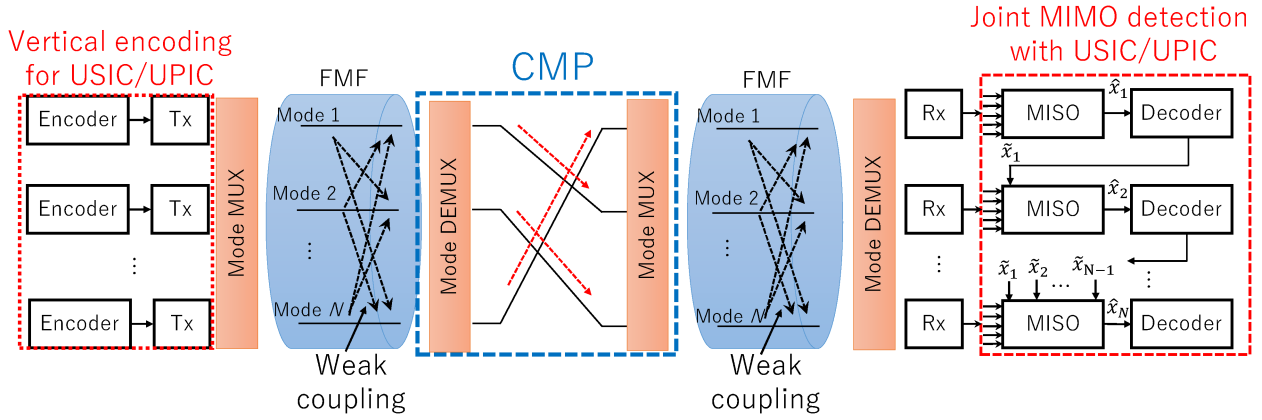


Fig. 2. Schematic of a MDM-MIMO system with CMP and USIC/UPIC schemes.

where we define  $\mathbf{x} \triangleq [x_1 \ x_2 \ \dots \ x_{N_T}]^T \in \mathbb{C}^{N_T \times 1}$ ,  $\mathbf{y} \triangleq [y_1 \ y_2 \ \dots \ y_{N_R}]^T \in \mathbb{C}^{N_R \times 1}$ , and  $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$  as the transmitted symbol vector, the received symbol vector, and the noise vector of the circularly symmetrical Gaussian with zero mean, respectively. The overall channel transfer matrix is also defined as  $\mathbf{H} \triangleq [\mathbf{h}_1 \ \mathbf{h}_2 \ \dots \ \mathbf{h}_{N_T}] \in \mathbb{C}^{N_R \times N_T}$  where  $\mathbf{h}_i \in \mathbb{C}^{N_R \times 1}$  denotes the  $i$ -th column vector of  $\mathbf{H}$ . At the receiver end, MIMO-DSP techniques are usually required to correctly extract data streams from highly-mixed MDM signals. Linear filtering detection techniques are attractive for application to MDM transmission because of their efficient computational implementation property. The main objective of linear detection is to obtain an estimate of an  $i$ -th transmitted signal  $\hat{x}_i$ , by multiplying the  $i$ -th column vector of the filter weight matrix  $\mathbf{W} \in \mathbb{C}^{N_R \times N_T}$  with  $\mathbf{y}$ :

$$\hat{x}_i = \mathbf{W}_{:,i}^H \mathbf{y}. \quad (2)$$

In what follows, a minimum mean square error (MMSE) filter is considered as one popular design criterion of  $\mathbf{W}$ . Then the product in (2) yields

$$\hat{x}_i = \mathbf{W}_{:,i}^H \mathbf{h}_i x_i + \sum_{j \neq i} \mathbf{W}_{:,i}^H \mathbf{h}_j x_j + \mathbf{W}_{:,i}^H \mathbf{n}. \quad (3)$$

In (3), the second term represents residual interference onto  $i$ -th stream from other streams. This intermodal interference is enhanced when  $\mathbf{H}$  exhibits a nonunitary property and hence MDL exists. In such a scenario, the result is degraded post-processing SINR, which has a negative impact on achievable MIMO capacity and transmission distance.

### B. Interference Cancellation Techniques

One possible solution to address aforementioned the issue is to introduce nonlinear detection techniques including successive/parallel interference cancellation (SIC/PIC). Technical challenges in applying these interference cancelers to optical MDM transmission include addressing error propagation and developing robust channel estimation method with low complexity. We earlier proposed a novel MIMO-DSP technique [11], [23] that was designed to mitigate MDL impact by improving SINR in long-haul MDM transmission. Significant features of

this technique are that it performs interference nulling and subtraction without estimating  $\mathbf{H}$  and generating replica signals. We developed and proposed two types of interference cancellers: successive [23] and parallel processing [11], although in the presented paper we focus on the former, which we term unreplicated SIC (USIC).

As shown in Fig. 2, we assume a transceiver where a single data stream is divided into parallel  $N_T$  streams that are independently encoded. After transmission, a USIC detector is used for data stream detection. It has a layered structure comprising multiple detection stages, each responsible for extracting a single stream from  $\mathbf{y}$ . In the initial detection stage, stream detection for  $x_1$  is performed by a linear detection scheme without any interference cancellation. Then a soft estimate of  $x_1$ , denoted as  $\tilde{x}_1$ , is constructed on the basis of reliability information of bits/symbols (e.g., log-likelihood ratios) delivered from a decoder, and used for interference cancellation at a second detection stage. Likewise, stream detection is successively performed using soft estimates already detected and decoded in the preceding stages. Note that, at a  $i$ -th detection stage, an augmented received signal vector  $\mathbf{y}^{(i)} \in \mathbb{C}^{(N_R+i-1) \times 1}$  is used as an input for a multiple-input single-output (MISO) detector, which is constructed by concatenating the received signal vector and obtained signal estimates in the preceding detection stages:

$$\mathbf{y}^{(i)} = \begin{bmatrix} \mathbf{y} \\ \tilde{x}_1 \\ \vdots \\ \tilde{x}_{i-1} \end{bmatrix}. \quad (4)$$

In USIC detection, filter weight vector used at a  $i$ -th stage  $\mathbf{w}^{(i)} \in \mathbb{C}^{(N_R+i-1) \times 1}$  is easily obtained by the stochastic gradient approach for linear filter adaptation techniques including the least mean square (LMS) algorithm. Consequently, the USIC receiver enables us to obtain the  $i$ -th signal estimate  $\hat{x}_i$ :

$$\hat{x}_i = \mathbf{w}^{(i)H} \mathbf{y}^{(i)}. \quad (5)$$

This operation realizes computationally efficient receivers because it omits the channel estimation processing that is usually required in conventional SIC receivers. It should be noted that the aforementioned transceiver/receiver in MDM-MIMO systems is

similar to the vertical Bell laboratories space-time (V-BLAST) architecture with a MMSE-SIC receiver [24], which is known as a system that attains open-loop MIMO capacity when combined with per-antenna rate control [21], [25].

A comparative study on the performance of MIMO detection schemes is presented in our previous works [11], which revealed that the detection performance provided by the proposed successive/parallel cancelers, each respectively corresponding to USIC and unreplicated parallel interference canceler (UPIC), is principally identical to that conventional SIC/PIC schemes. We again emphasize that the main advantage of the proposed schemes is a reduced computational cost: for USIC detection, the required complexity order for detection of  $N$  data streams becomes  $O(N^2)$  which is the same order with conventional MIMO equalizers. UPIC detection generally outperforms USIC detection at the expense of enhanced computational complexity proportional to the number of outer iterations  $M$  (i.e.,  $O(MN^2)$ ).

#### IV. PROJECTION TECHNIQUES FOR SIGNAL DETECTION IN MDM-MIMO SYSTEMS

The aim of this section is to discuss practical design criteria of the filter weight vector  $\mathbf{w}$  for applications to conventional MIMO equalization or USIC/UPIC techniques. Furthermore we formulate a detection scheme based on a projection approach which can be directly applicable for MDM-MIMO systems, termed MIMO-APA. To this end, we consider a more general case of the MIMO model to have a time-dispersion effect with a memory length of  $L$ . Accordingly, the MIMO-based receiver size is expanded with  $\mathbf{y} \in \mathbb{C}^{N_R L \times 1}$  and  $\mathbf{w} \in \mathbb{C}^{N_R L \times 1}$ .

##### A. Linear Filter Adaptation Techniques

An optimum filter in the mean square error (MSE) sense is designed with a quadratic cost function evaluating a difference between a desired signal and a filter output, known as a MMSE filter. A stochastic and practical approach to attain a MMSE solution is the LMS algorithm that updates a filter weight vector  $\mathbf{w}$  on the basis of an instantaneous estimate of MSE. In fact, MIMO equalization with LMS-based weight updating has considerably contributed to recently-achieved high-capacity and/or long-haul MDM transmission experiments. Learning curve behavior of LMS technique is, however, governed by eigenvalue spread of a covariance matrix of a received vector  $\mathbf{R}$  [26]. This fact indicates that LMS convergence property may be degraded in the presence of MDL.

An alternative criterion in designing linear filters is derived from a deterministic approach that minimizes the (weighted) squared error expectation, known as the recursive least squares (RLS) algorithm. In RLS adaptation, filter convergence is less sensitive to eigenvalue spread of  $\mathbf{R}$ , suggesting advantages for application in MDM transmission with large MDL [27]. The main drawbacks of RLS include numerical instability and high computational complexity on the order of  $O(N_R^2 L^2)$ . A compromise solution that achieves relatively fast convergence for  $\mathbf{w}$  with moderate complexity is provided by the affine projection algorithm (APA) [28]. At sample time  $k$ , APA updates  $\mathbf{w}(k)$  with an adjustment vector  $\Delta\mathbf{w}(k)$  using affine projection onto

subspace spanned by the current input vector along with previous  $p - 1$  input vectors, where  $p$  is the projection order. As will be shown in Section IV-B, the required complexity for the APA is at least on the order of  $O(N_R p^2 L)$ . Note that APA is considered as a generalization of the well-known normalized LMS (NLMS) algorithm (i.e., a vector operation version of APA with  $p = 1$ ).

##### B. Affine Projection for MDM-MIMO Systems: MIMO-APA

To construct a projection technique designed for MIMO systems, we start with a derivation of MIMO-APA. To make discussions simple, we focus on signal detection for a single data stream (i.e., MISO detection), because a linear MIMO detector can be substantially decomposed into a parallelized set of MISO detectors. Also assumed is the transmission regime with small  $p$  relative to  $L$  (i.e.,  $p \ll L$ ), because considered dispersive channel mainly comes from the presence of DMD which generally causes spreading of the impulse response width as wide as several tens of ns, as stated in Section II.

We first divide a filter weight vector  $\mathbf{w}(k)$  into filter weight sub-vectors  $\mathbf{w}_i(k) \in \mathbb{C}^{L \times 1}$ , and a received signal vector  $\mathbf{y}(k)$  into received signal sub-vectors  $\mathbf{y}_i(k) \in \mathbb{C}^{L \times 1}$ , each obtained from  $N_R$  receivers. These quantities are related as

$$\mathbf{w}(k) = \begin{bmatrix} \mathbf{w}_1(k) \\ \mathbf{w}_2(k) \\ \vdots \\ \mathbf{w}_{N_R}(k) \end{bmatrix}, \quad (6)$$

and

$$\mathbf{y}(k) = \begin{bmatrix} \mathbf{y}_1(k) \\ \mathbf{y}_2(k) \\ \vdots \\ \mathbf{y}_{N_R}(k) \end{bmatrix}. \quad (7)$$

We also define an  $i$ -th received signal sub-matrix  $\mathbf{Y}_i(k) \in \mathbb{C}^{L \times p}$  and a desired signal vector  $\mathbf{d}(k) \in \mathbb{C}^{1 \times p}$ , respectively expressed as

$$\begin{aligned} \mathbf{Y}_i(k) &\triangleq [\mathbf{y}_i(k) \quad \mathbf{y}_i(k-1) \quad \cdots \quad \mathbf{y}_i(k-p+1)] \\ &= \begin{bmatrix} y_i(k) & y_i(k-1) & \cdots & y_i(k-p+1) \\ y_i(k-1) & y_i(k-2) & \cdots & y_i(k-p) \\ \vdots & \vdots & \ddots & \vdots \\ y_i(k-L+1) & y_i(k-L) & \cdots & y_i(k-p-L+2) \end{bmatrix}, \end{aligned} \quad (8)$$

and

$$\mathbf{d}(k) \triangleq [d(k) \quad d(k-1) \quad \cdots \quad d(k-p+1)]. \quad (9)$$

We consider a minimization problem for the set of the squared Euclidean norm change in  $\mathbf{w}_i(k)$ , subject to the constraint of

$$\sum_{i=1}^{N_R} \mathbf{w}_i^H(k) \mathbf{Y}_i(k) = \mathbf{d}(k). \quad (10)$$



TABLE I  
ALGORITHM 1: MIMO-APA

| Line# | Equation   | Complexity      |
|-------|--|-----------------|
| 1:    | <b>for</b> $i = 1$ to $N_R$ <b>do</b>  |                 |
| 2:    | $\mathbf{Y}_i(k) = [\mathbf{y}_i(k) \mathbf{Y}_{i :,1:p-1}(k-1)]$                      | -               |
| 3:    | <b>end for</b>   |                 |
| 4:    | $\mathbf{e}(k) = \mathbf{d}(k) - \sum_{i=1}^{N_R} \mathbf{w}_i^H(k-1) \mathbf{Y}_i(k)$ | $N_R p L$       |
| 5:    | $\mathbf{R}(k) = \sum_{i=1}^{N_R} \mathbf{Y}_i^H(k) \mathbf{Y}_i(k)$                   | $N_R p^2 L$     |
| 6:    | $\mathbf{R}_{\text{inv}}(k) = \mathbf{R}^{-1}(k)$                                      | $p^3/2 + p^2/2$ |
| 7:    | $\mathbf{g}(k) = \mathbf{R}_{\text{inv}}(k) \mathbf{e}^H(k)$                           | $p^2$           |
| 8:    | $\mathbf{g}'(k) = \mu \mathbf{g}(k)$   | $p$             |
| 9:    | <b>for</b> $i = 1$ to $N_R$ <b>do</b>  |                 |
| 10:   | $\Delta \mathbf{w}_i(k) = \mathbf{Y}_i(k) \mathbf{g}'(k)$                              | $N_R p L$       |
| 11:   | $\mathbf{w}_i(k) = \mathbf{w}_i(k-1) + \Delta \mathbf{w}_i(k)$                         | -               |
| 12:   | <b>end for</b>   |                 |

This is solved with the aid of the method of Lagrange multipliers where a cost function  $J(k)$  is defined with Lagrange multipliers  $\lambda_j$  ( $1 \leq j \leq p$ ) [28]:

$$J(k) \triangleq \sum_{i=1}^{N_R} \|\mathbf{w}_i(k) - \mathbf{w}_i(k-1)\|^2 + \sum_{j=1}^p \text{Re} \left[ \lambda_j^* \left[ d(k-j+1) - \sum_{i=1}^{N_R} \mathbf{w}_i^H(k) \mathbf{y}_i(k-j+1) \right] \right]. \quad (11)$$

Differentiating (11) with respect to each  $\mathbf{w}_i^*(k)$  gives a collective form of MIMO-APA whose derivation is provided in Appendix:

$$\hat{x}(k) = \sum_{i=1}^{N_R} \mathbf{w}_i^H(k-1) \mathbf{y}_i(k), \quad (12)$$

$$\mathbf{e}(k) \triangleq \mathbf{d}(k) - \sum_{i=1}^{N_R} \mathbf{w}_i^H(k-1) \mathbf{Y}_i(k), \quad (13)$$

$$\Delta \mathbf{w}_i(k) \triangleq \mu \mathbf{Y}_i(k) \mathbf{g}(k), \quad (14)$$

$$\mathbf{w}_i(k) = \mathbf{w}_i(k-1) + \Delta \mathbf{w}_i(k), \quad (15)$$

where we define a step-size parameter  $\mu$ , a priori error vector  $\mathbf{e}(k) \in \mathbb{C}^{1 \times p}$ , an adjustment sub-vector  $\Delta \mathbf{w}_i(k) \in \mathbb{C}^{L \times 1}$ , and a *pre-filtering vector*  $\mathbf{g}(k) \in \mathbb{C}^{p \times 1}$ .  $\mathbf{g}(k)$  is calculated by a correlation sub-matrix  $\mathbf{R}_i(k) \in \mathbb{C}^{p \times p}$ , and a correlation matrix  $\mathbf{R}(k) \in \mathbb{C}^{p \times p}$ , each defined as

$$\mathbf{R}_i(k) \triangleq \mathbf{Y}_i^H(k) \mathbf{Y}_i(k), \quad (16)$$

$$\mathbf{R}(k) \triangleq \sum_{i=1}^{N_R} \mathbf{R}_i(k), \quad (17)$$

and

$$\mathbf{g}(k) \triangleq \mathbf{R}^{-1}(k) \mathbf{e}^H(k). \quad (18)$$

Table I summarizes the MIMO-APA together with a number of required complex multiplications/divisions (i.e., computational complexity). Note that we use the fact that the inverse computation for an invertible matrix  $\mathbf{A} \in \mathbb{C}^{N \times N}$  requires  $\frac{N^3}{2} + \frac{N^2}{2}$  complex multiplications or divisions based on the Gauss-Jordan elimination algorithm. We noticed that the majority of the computational efforts in MIMO-APA is governed in computation

steps for calculation of a priori error vector  $\mathbf{e}(k)$  (line#4), a correlation matrix  $\mathbf{R}(k)$  (line#5), and an adjustment sub-vector  $\Delta \mathbf{w}_i(k)$  (lines#9 to #11). One negative feature of MIMO-APA is that the required complexity grows when the projection order  $p$  is increased, indicating the tradeoff between a fast convergence and a computational cost.

### C. Proposed: Fast MIMO Affine Projection Algorithm (Fast MIMO-APA)

This subsection is devoted to an introduction of a computationally efficient version of the MIMO-APA with much decreased complexity, which we call ‘‘Fast MIMO-APA’’ in the presented paper. The complexity reduction is possible if we introduce some auxiliary vectors by which we can make full use of the *time-shifting* property of variables to omit redundant computation at succeeding processing intervals. This is achieved by dividing a weight vector  $\mathbf{w}$  into sub-weight vectors  $\mathbf{w}_i$  as we have done in Section IV-B. Note that the presented approach is essentially identical to that was originally proposed in [29] for echo canceling systems, while we expand its application from single-input single-output (SISO) systems to MIMO systems.

A priori error calculation in MIMO-APA is simplified by using the constraint at  $k-1$  of

$$\sum_{i=1}^{N_R} \mathbf{w}_i^H(k-1) \mathbf{Y}_i(k-1) = \mathbf{d}(k-1), \quad (19)$$

which gives an equivalent form of (13) as [29]

$$\mathbf{e}(k) = [d(k) - \hat{x}(k) \quad (1-\mu)e_{1:p-1}(k-1)]. \quad (20)$$

Equation (20) indicates that  $\mathbf{e}(k)$  is recursively obtained only with  $p-1$  complex multiplications.

To obtain a recursive form of a correlation sub-matrix  $\mathbf{R}_i$  for its updating, we introduce a correlation sub-vector  $\mathbf{r}_i(k) \in \mathbb{C}^{1 \times p}$  defined as

$$\mathbf{r}_i(k) \triangleq \mathbf{y}_i^H(k) \mathbf{Y}_i(k). \quad (21)$$

Considering a time-shifting property between  $\mathbf{Y}_i(k)$  and  $\mathbf{Y}_i(k-1)$ , an updating equation of  $\mathbf{r}_i(k)$  is obtained as

$$\mathbf{r}_i(k) = \mathbf{r}_i(k-1) + y_i^*(k) \mathbf{Y}_{i|1,:}(k) - y_i^*(k-L) \mathbf{Y}_{i|k-L+1,:}(k-1). \quad (22)$$

The fact that  $\mathbf{R}_i(k)$  comprises  $\mathbf{r}_i(k)$  and  $\mathbf{R}_i(k-1)$  leads a recursive updating form of

$$\mathbf{R}_i(k) = \begin{pmatrix} \mathbf{r}_{i|1}(k) & \mathbf{r}_{i|2:p}(k) \\ \mathbf{r}_{i|2:p}^H(k) & \mathbf{R}_{i|1:p-1,1:p-1}(k-1) \end{pmatrix}. \quad (23)$$

From the results of (17) and (23), the update of  $\mathbf{R}(k)$  requires only  $2N_R p$  complex multiplications associated with the calculation of  $\mathbf{r}_i(k)$ .

Next we consider an alternative of a filter weight sub-vector  $\mathbf{w}_i(k)$ . We rewrite (14) as

$$\Delta \mathbf{w}_i(k) = \sum_{j=1}^p \mu \mathbf{g}_{|j}(k) \mathbf{y}_i(k-j+1). \quad (24)$$

Equations (15) and (24) indicate that the weight sub-vector at time  $k$  comprises a purely linear combination of the input sub-vectors from the past to the current. Therefore a *smoothing coefficient*  $s_{(j)}(k)$  is introduced as the coefficient of  $\mathbf{y}_i(k-j+1)$  at time  $k$  that satisfies a relation of

$$\mathbf{w}_i(k) = \sum_{j=1}^k s_{(j)}(k) \mathbf{y}_i(k-j+1). \quad (25)$$

To obtain an updating equation of  $s_{(j)}(k)$ , (15) is reformed as

$$\begin{aligned} \mathbf{w}_i(k) &= \mathbf{w}_i(k-1) + \Delta \mathbf{w}_i(k) \\ &= \sum_{j=1}^k s_{(j)}(k-1) \mathbf{y}_i(k-j) + \sum_{j=1}^p \mu \mathbf{g}_{|j}(k) \mathbf{y}_i(k-j+1). \end{aligned} \quad (26)$$

Comparison of (25) with (26) directly gives updating equations of a smoothing coefficient  $s_{(j)}(k)$ :

$$s_{(j)}(k) = \begin{cases} \mu \mathbf{g}_{|j}(k) & (j=1) \\ s_{(j-1)}(k-1) + \mu \mathbf{g}_{|j}(k) & (2 \leq j \leq p) \\ s_{(j)}(k-1) & (p+1 \leq j \leq k) \end{cases} \quad (27)$$

By introducing a smoothing coefficient vector  $\mathbf{s}(k) \triangleq [s_{(1)}(k) \ s_{(2)}(k) \ \cdots \ s_{(p)}(k)]^T \in \mathbb{C}^{p \times 1}$ , more compact form of (27) is simply expressed as

$$\mathbf{s}(k) = \begin{bmatrix} 0 \\ \mathbf{s}_{|1:p-1}(k-1) \end{bmatrix} + \mu \mathbf{g}(k). \quad (28)$$

We further divide (25) into ones containing  $s_{(j)}(k)$  that are updated at time  $k$  and the rest:

$$\begin{aligned} \mathbf{w}_i(k) &= \sum_{j=1}^k s_{(j)}(k) \mathbf{y}_i(k-j+1) \\ &= \sum_{j=1}^{p-1} s_{(j)}(k) \mathbf{y}_i(k-j+1) + \sum_{j=p}^k s_{(j)}(k) \mathbf{y}_i(k-j+1) \\ &= \sum_{j=1}^{p-1} s_{(j)}(k) \mathbf{y}_i(k-j+1) + \mathbf{z}_i(k), \end{aligned} \quad (29)$$

where we use an *approximation vector*  $\mathbf{z}_i(k) \in \mathbb{C}^{L \times 1}$  defined as

$$\mathbf{z}_i(k) \triangleq \sum_{j=p}^k s_{(j)}(k) \mathbf{y}_i(k-j+1). \quad (30)$$

From its definition, an updating equation for  $\mathbf{z}_i(k)$  becomes

$$\mathbf{z}_i(k) = \mathbf{z}_i(k-1) + s_{(p)}(k) \mathbf{y}_i(k-p+1). \quad (31)$$

Finally, by evaluating (29) at  $k-1$ , an alternative form for signal estimate calculation is now obtained by multiplying  $\mathbf{y}_i$  with  $\mathbf{w}_i(k-1)$  both sides:

$$\hat{\mathbf{x}}(k) = \sum_{i=1}^{N_R} [\mathbf{y}_i^H(k) \mathbf{z}_i(k-1) + \mathbf{r}_{i|1:p-1}(k) \mathbf{s}_{|1:p-1}(k-1)]^*. \quad (32)$$

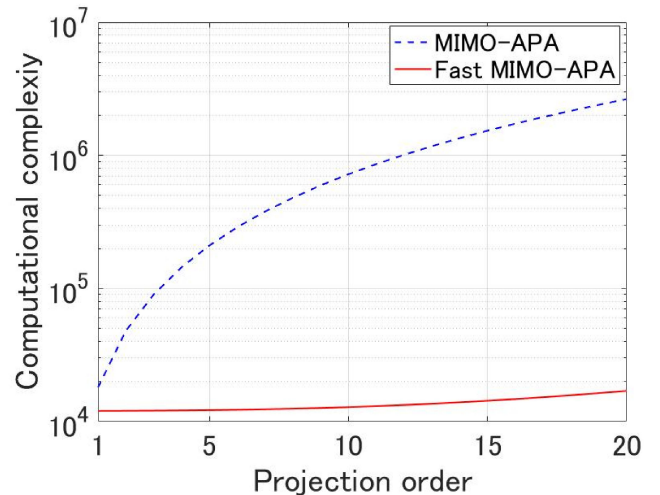


Fig. 3. Comparison of the required complexity in MIMO-APA and Fast MIMO-APA as a function of the projection order  $p$  when  $L = 1000$  and  $N_R = 6$ .

The important feature in (32) is that  $\hat{\mathbf{x}}(k)$  is obtained without a direct use of  $\mathbf{w}_i(k)$ . This means that the use of  $\mathbf{z}_i(k)$  instead of  $\mathbf{w}_i(k)$  enables us to avoid a heavy computational cost in its updating as required in (14).

The processing flow of a Fast MIMO-APA technique is summarized in Table II. Most importantly, the major part of computational complexity no longer relies on the projection order  $p$  in the regime of  $p \ll L$ , whereas signal detection performance of a Fast MIMO-APA is strictly preserved relative to that of a MIMO-APA. Although not the scope of the presented paper, in the regime of  $p$  approaching  $L$ , the computation of  $\mathbf{R}^{-1}(k)$  and  $\mathbf{g}(k)$  dominantly impacts on the overall complexity. This can be mitigated from  $O(p^3)$  to  $O(p)$  based on the approach of the fast transversal filter (FTF) algorithm [29], [30], although it still involves with numerical instability problems.

At the final part of this section, we provide a comparative example of the complexity required both for MIMO-APA and Fast MIMO-APA. To do this, we set the equalizer memory length of the channel  $L$  to 1000 taps and the number of receivers  $N_R$  to 6. This assumes that 10-Gbaud MDM PDM signals are processed by linear MIMO detectors operating at 2 samples per symbol after transmission over 2-LP FMFs which causes spreading of the impulse response width with a 50-ns window. Fig. 3 compares the required complexity per single symbol output for both schemes. For a detection using MIMO-APA, the complexity cost significantly grows with the increased number of  $p$ . On the other hand, if we apply the proposed Fast MIMO-APA, much smaller complexity is required because of lower dependency on  $p$ . In particular, a complexity of Fast MIMO-APA at  $p = 10$  is reduced by a factor of 0.02 in comparison to the use of MIMO-APA.

## V. EXPERIMENTAL SETUP AND RESULTS

To examine convergence behavior of LMS and APA in the presence of MDL, a transmission experiment was conducted

TABLE II  
ALGORITHM 2: FAST MIMO-APA

| Line# | Equation  | Complexity      |
|-------|---|-----------------|
| 1:    | <b>for</b> $i = 1$ to $N_R$ <b>do</b>   |                 |
| 2:    | $\mathbf{Y}_i(k) = [\mathbf{y}_i(k) \mathbf{Y}_{i 1:p-1}(k-1)]$   | -               |
| 3:    | $\mathbf{r}_i(k) = \mathbf{r}_i(k-1) + \mathbf{y}_i^*(k)\mathbf{Y}_{i 1:p-1}(k) - \mathbf{y}_i^*(k-L)\mathbf{Y}_{i k-L+1:p-1}(k-1)$ | $2N_R p$        |
| 4:    | <b>end for</b>  |                 |
| 5:    | update $\mathbf{R}(k)$ , and compute $\mathbf{R}_{\text{inv}}(k) = \mathbf{R}^{-1}(k)$  | $p^3/2 + p^2/2$ |
| 6:    | $\hat{\mathbf{x}}(k) = \sum_{i=1}^{N_R} [\mathbf{y}_i^H(k)\mathbf{z}_i(k-1) + \mathbf{r}_{i 1:p-1}(k)\mathbf{s}_{i 1:p-1}(k-1)]^*$  | $N_R(L+p-1)$    |
| 7:    | $\mathbf{e}(k) = [d(k) - \hat{\mathbf{x}}(k) \quad (1-\mu)\mathbf{e}_{i 1:p-1}(k-1)]$   | $p-1$           |
| 8:    | $\mathbf{g}(k) = \mathbf{R}_{\text{inv}}(k)\mathbf{e}^H(k)$   | $p^2$           |
| 9:    | $\mathbf{g}'(k) = \mu\mathbf{g}(k)$   | $p$             |
| 10:   | $\mathbf{s}(k) = \begin{bmatrix} 0 \\ \mathbf{s}_{i 1:p-1}(k-1) \end{bmatrix} + \mathbf{g}'(k)$                                     | -               |
| 11:   | <b>for</b> $i = 1$ to $N_R$ <b>do</b>   |                 |
| 12:   | $\mathbf{z}_i(k) = \mathbf{z}_i(k-1) + \mathbf{s}_{i p}(k)\mathbf{x}_i(k-p+1)$  | $N_R L$         |
| 13:   | <b>end for</b>  |                 |

where 3-mode MDM PDM signals were transmitted with three-fold recirculating-loop systems comprising 12-core  $\times$  3-mode MC-FMF with 52.7-km length. Each carrier comprised dual subcarriers, driven at 6 Gbaud. The transmission frame was the 33040-symbol-length QPSK-modulated pattern. We previously provided a more detailed description of the transmission setup [10]. At the receiver, an offline processing was performed including the chromatic dispersion compensation, the USIC detection, and the digital PLL-based carrier phase recovery [31], where  $w_i$  is updated by the time-domain  $6 \times 6$  MIMO processing based either on the LMS or Fast MIMO-APA with  $L = 400$ . In our DSP steps, the frequency offset processing was not included, because we employed stable external cavity tunable lasers as laser sources in the conducted experiments. Since the convergence performance is considered to be affected by the equalizer operation mode employed in each filter adaptation scheme, we performed all the equalization by the data-aided mode for a fair comparison purpose. Frame synchronization was performed using correlations between repetitive patterns [32]. A perfect symbol reconstruction by the decoders (i.e., no symbol error) was assumed for all streams used in the USIC detection scheme for evaluation simplicity, and RLS algorithm was not used due to huge computational complexity.

Prior to representing the experimental evaluation results, it is worthwhile to discuss the computational complexity for the LMS based and the Fast MIMO-APA based adaptations, when we apply the USIC scheme for data stream detection. In the LMS-based filter adaptation, the processing at the  $i$ -th USIC detection contains  $(N+i-1)L$  complex multiplications for the computation of the filter output, and  $(N+i-1)L+1$  for the update of the filter weight vector [11]. As mentioned in IV-C, in the filter adaptation based on proposed Fast MIMO-APA at the  $i$ -th USIC detection, we found that the majority of the computational complexity also comes from the computation of the filter output (line# 6 in Table II) and the update of filter weight vector (line# 12 in Table II), each approximately requiring  $(N+i-1)L$  and  $(N+i-1)L$  complex multiplications, respectively, in the regime of  $p \ll L$ . This means that, for the practical application of these schemes for MDM-MIMO transmission, the computational complexity for both adaptation schemes becomes almost identical at each USIC detection stage.

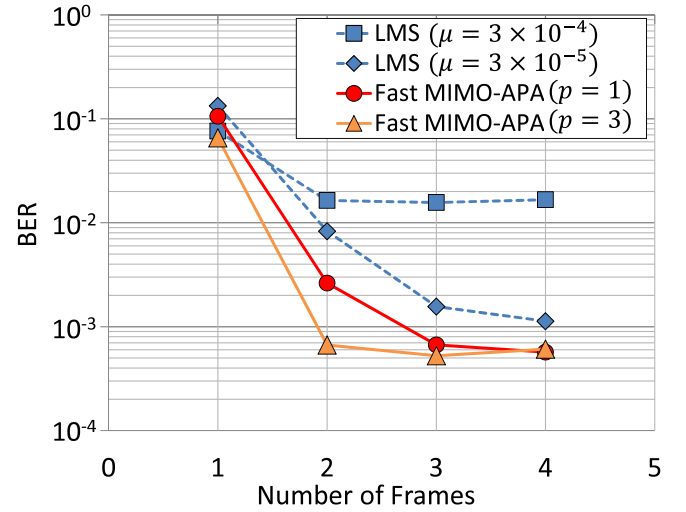


Fig. 4. BER convergence curves in the detection of mode 1 after transmission over 2500-km MC-FMF for core#4 obtained by USIC detection based on LMS or Fast MIMO-APA.

Consequently, denoting  $C_1$  and  $C_2$  as the overall complexity in detecting  $N$  data streams by the USIC scheme based on the LMS or Fast MIMO-APA, respectively, they are evaluated as

$$C_1 = C_2 \approx \sum_{i=1}^N 2(N+i-1)L \quad (33)$$

$$= N(3N-N)L. \quad (34)$$

Fig. 4 represents the BER curves obtained for mode 1 after 2500-km transmission over core#4 where a large MDL was observed exceeding 10 dB [10]. In the following evaluation, we focus on two properties characterizing the adaptive filter performance: the convergence speed and the excess MSE. The convergence speed is a decreasing rate in the learning curve, appeared as the transient performance of the cost function which is defined as the  $L_2$ -norm difference between the desired signal and the filter output. The excess MSE is a metric that quantifies the residual error relative to that can be attained by the optimal Wiener filter. In the adaptive filter theory, it is a well-known fact

that the step-size parameter for the LMS adaptation, denoted as  $\mu_1$ , governs the convergence speed. If  $\mu_1$  is chosen satisfying  $0 < \mu_1 < \frac{2}{\text{trace}(\mathbf{R})}$ , the steady-state excess MSE is shown to scale with  $\mu$  [26], meaning that there exists the tradeoff between the fast convergence property and the amount of excess MSE. A larger step-size parameter of the Fast MIMO-APA, denoted as  $\mu_2$ , also accelerates the convergence speed, although the excess MSE increases with both  $\mu_2$  and  $p$  [33]. In the presented work, we set  $\mu_2$  to  $10^{-2}$  to see a learning curve behavior dependence on  $p$ . For LMS adaptation with  $\mu$  of  $3 \times 10^{-5}$ , the BER curve did not converge even using four frames (i.e., 0.13 M symbols). When larger  $\mu$  of  $3 \times 10^{-4}$  was employed, a faster convergence was obtained at the sacrifice of BER performance. This might be due to the increased excess MSE of a filter weight vector. The convergence speed was greatly improved by the Fast MIMO-APA with the increased projection order  $p$ . These results indicate that the use of the USIC scheme together with the Fast MIMO-APA enhances the convergence property of detection schemes in MDL-impaired systems, enabling spectrally efficient MDM transmission with a decreased overhead portion in training sequences for optical MDM channels with fast variations which require periodic insertion of training sequences for a stable channel tracking [34].

## VI. CONCLUSION

We have reviewed issues to be addressed in achieving long-haul high-capacity SDM transmission in MDM-MIMO systems. We presented key enablers to enhance MDM transmission performance in the presence of DMD and MDL, the including CMP transmission scheme and USIC/UPIC detection scheme. As a practical issue in MDM transmission, we also discussed adaptation schemes for MDM-MIMO signal detection using linear filtering techniques. A novel reduced-complexity projection technique, termed Fast MIMO-APA, was newly proposed and shown that a convergence performance was improved by adopting the proposed method in conjunction with an SIC-based receiver.

## APPENDIX

We derive a formulation of MIMO-APA described in Subsection IV-B. We first take a derivative of  $J(k)$  in (11) with respect to each  $\mathbf{w}_i^*(k)$ :

$$\frac{\partial J(k)}{\partial \mathbf{w}_i^*(k)} = (\mathbf{w}_i(k) - \mathbf{w}_i(k-1)) - \mathbf{Y}_i(k)\boldsymbol{\lambda}^*, \quad (35)$$

where a Lagrange multiplier vector  $\boldsymbol{\lambda} \in \mathbb{C}^{p \times 1}$  is defined as  $\boldsymbol{\lambda} \triangleq [\lambda_1 \cdots \lambda_p]^T$ . By setting (35) to be zero, an updating equation of  $\mathbf{w}_i(k)$  becomes

$$\mathbf{w}_i(k) = \mathbf{w}_i(k-1) + \mathbf{Y}_i(k)\boldsymbol{\lambda}^*. \quad (36)$$

Substituting (36) into (10) gives

$$\mathbf{d}(k) - \sum_{i=1}^{N_R} [\mathbf{w}_i(k-1) + \mathbf{Y}_i(k)\boldsymbol{\lambda}^*]^H \mathbf{Y}_i(k) = 0. \quad (37)$$

After some algebraic manipulation, (37) is simplified as

$$\mathbf{e}(k) = \boldsymbol{\lambda}^T \sum_{i=1}^{N_R} \mathbf{R}_i(k) \quad (38)$$

$$= \boldsymbol{\lambda}^T \mathbf{R}(k), \quad (39)$$

where  $\mathbf{e}(k)$ ,  $\mathbf{R}_i(k)$ , and  $\mathbf{R}(k)$  are defined in (13), (16), and (17), respectively. Since we assume a full-column rank  $\mathbf{Y}_i(k)$  for all  $i$ ,  $\boldsymbol{\lambda}$  is obtained from (39):

$$\boldsymbol{\lambda} = \mathbf{R}^{-1}(k)\mathbf{e}^T(k). \quad (40)$$

Using (36) and (40) with an introduction of  $\mu$ , we obtain an updating equation of MIMO-APA for  $\mathbf{w}_i(k)$ :

$$\mathbf{w}_i(k) = \mathbf{w}_i(k-1) + \mu \mathbf{Y}_i(k)\mathbf{R}^{-1}(k)\mathbf{e}^H(k) \quad (41)$$

$$= \mathbf{w}_i(k-1) + \mu \mathbf{Y}_i(k)\mathbf{g}(k), \quad (42)$$

which coincides with (14), (15) and (18).

## ACKNOWLEDGMENT

Part of this research utilized results from research commissioned by the National Institute of Information and Communications Technology (NICT) of Japan.

## REFERENCES

- [1] V. Sleiffer *et al.*, "73.7 Tb/s ( $96 \times 3 \times 256$ -Gb/s) mode-division-multiplexed DP-16QAM transmission with inline MM-EDFA," *Opt. Express*, vol. 20, no. 26, pp. B428–B438, Nov. 2012.
- [2] E. Ip *et al.*, "146 $\lambda \times 6 \times 19$ -Gbaud wavelength-and mode-division multiplexed transmission over  $10 \times 50$ -km spans of few-mode fiber with a gain-equalized few-mode EDFA," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 790–797, Nov. 2014.
- [3] R. Ryf *et al.*, "Photonic-lantern-based mode multiplexers for few-mode-fiber transmission," in *Proc. 39th Opt. Fiber Commun. Conf. Exhib.*, San Francisco, CA, USA, Mar. 2014, Paper W4J.2.
- [4] R. Ryf *et al.*, "Distributed Raman amplification based transmission over 1050-km few-mode fiber," in *Proc. 41st Eur. Conf. Opt. Commun.*, Valencia, Spain, Sep. 2015, Paper Tu.3.2.3.
- [5] R. Ryf *et al.*, "72-tb/s transmission over 179-km all-fiber 6-mode span with two cladding pumped in-line amplifiers," in *Proc. Eur. Conf. Opt. Commun.*, Valencia, Spain, Sep. 2015, Paper Tu.3.2.2.
- [6] J. van Weerdenburg *et al.*, "138 Tbit/s transmission over 650 km graded-index 6-mode fiber," in *Proc. 43rd Eur. Conf. Exhib. Opt. Commun.*, Gothenburg, Sweden, Sep. 2017, Paper Th.PDPA.4.
- [7] G. Rademacher *et al.*, "159 Tbit/s C+L band transmission over 1045 km 3-mode graded-index few-mode fiber," in *Proc. 43rd Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, Mar. 2018, Paper Th4C.4.
- [8] K. Shibahara *et al.*, "Full c-band 3060-km dmd-unmanaged 3-mode transmission with 40.2-tb/s capacity using cyclic mode permutation," in *Proc. 44th Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, Mar. 2019, Paper W3F.2.
- [9] K. Shibahara *et al.*, "Dense SDM (12-core  $\times$  3-mode) transmission over 527 km with 33.2-ns mode-dispersion employing low-complexity parallel MIMO frequency-domain equalization," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 196–204, Jan. 2016.
- [10] K. Shibahara *et al.*, "DMD-unmanaged long-haul SDM transmission over 2500-km 12-core  $\times$  3-mode MC-FMF and 6300-km 3-mode fmf employing intermodal interference cancelling technique," in *Proc. 43rd Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, Mar. 2018, Paper Th4C.6.
- [11] K. Shibahara *et al.*, "Iterative unreplicated parallel interference canceler for MDL-tolerant dense SDM (12-core  $\times$  3-mode) transmission over 3000 km," in *Proc. 44th Eur. Conf. Exhib. Opt. Commun.*, Roma, Italy, Sep. 2018, Paper Tu3F.6.
- [12] R. Ryf *et al.*, "10-mode mode-multiplexed transmission over 125-km single-span multimode fiber," in *Proc. 41st Eur. Conf. Opt. Commun.*, Valencia, Spain, Sep. 2015, Paper PDP 3.3.



- [13] D. Soma, S. Beppu, Y. Wakayama, Y. Kawaguchi, K. Igarashi, and T. Tsuritani, "257-Tbit/s partial mimo-based 10-mode C+L-band wdm transmission over 48-km FMF," in *Proc. 43rd Eur. Conf. Exhib. Opt. Commun.*, Gothenburg, Sweden, Sep. 2017, Paper M2.E.3.
- [14] G. Rademacher *et al.*, "93.34 Tbit/s/mode (280 Tbit/s) transmission in a 3-mode graded-index few-mode fiber," in *Proc. 43rd Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, Mar. 2018, Paper W4C.3.
- [15] R. Ryf *et al.*, "High-spectral-efficiency mode-multiplexed transmission over graded-index multimode fiber," in *Proc. 44th Eur. Conf. Exhib. Opt. Commun.*, Roma, Italy, Sep. 2018, Paper Th3B.1.
- [16] G. Rademacher *et al.*, "3500-km mode-multiplexed transmission through a three-mode graded-index few-mode fiber link," in *Proc. 43rd Eur. Conf. Opt. Commun.*, Gothenburg, Sweden, Sep. 2017, Paper M.2.E.4.
- [17] J. van Weerdenburg *et al.*, "Mode-multiplexed 16-QAM transmission over 2400-km large-effective-area depressed-cladding 3-mode fiber," in *Proc. 43rd Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, Mar. 2018, Paper W4C.2.
- [18] K. Shibahara, T. Mizuno, and Y. Miyamoto, "Long-haul mode multiplexing transmission enhanced by interference cancellation technique," in *Proc. 45th Eur. Conf. Exhib. Opt. Commun.*, Dublin, Ireland, Sep. 2019, Paper W.2.A.1.
- [19] J. M. Kahn, K.-P. Ho, and M. B. Shemirani, "Mode coupling effects in multi-mode fibers," in *Proc. 37th Opt. Fiber Commun. Conf. Exhib.*, Los Angeles, CA, USA, Mar. 2012, Paper OW3D.3.
- [20] Y. Sasaki, K. Takenaga, S. Matsuo, K. Aikawa, and K. Saitoh, "Few-mode multicore fibers for long-haul transmission line," *Opt. Fiber Technol.*, vol. 35, pp. 19–27, 2017.
- [21] P. J. Winzer and G. J. Foschini, "Mimo capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Express*, vol. 19, pp. 16680–16696, Aug. 2011.
- [22] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Opt. Express*, vol. 19, pp. 16612–16635, Aug. 2011.
- [23] K. Shibahara, T. Mizuno, and Y. Miyamoto, "LDPC-coded FMF transmission employing unreplicated successive interference cancellation for MDL-impact mitigation," in *Proc. 43rd Eur. Conf. Opt. Commun.*, Gothenburg, Sweden, Sep. 2017, Paper Th.1.D.4.
- [24] P. W. Wolniansky, G. J. Foschini, G. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals, Syst., Electron.*, 1998, pp. 295–300.
- [25] S. T. Chung, A. Lozano, H. C. Huang, A. Sutivong, and J. M. Cioffi, "Approaching the MIMO capacity with a low-rate feedback channel in V-BLAST," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 5, pp. 762–771, May 2004.
- [26] S. S. Haykin, *Adaptive Filter Theory*, Pearson Education India, New Delhi, 2008.
- [27] S. O. Arik, D. Askarov, and J. M. Kahn, "Adaptive frequency-domain equalization in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 32, pp. 1841–1852, May 2014.
- [28] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electron. Commun. Jpn. (Part I: Commun.)*, vol. 67, pp. 19–27, May 1984.
- [29] M. Tanaka, Y. Kaneda, S. Makino, and J. Kojima, "Fast projection algorithm for adaptive filtering," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E78-A, no. 10, pp. 1355–1361, Oct. 1995.
- [30] J. Cioffi and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 304–337, Apr. 1984.
- [31] T. Kobayashi *et al.*, "160-gb/s polarization-multiplexed 16-QAM long-haul transmission over 3,123 km using digital coherent receiver with digital PLL based frequency offset compensator," in *Proc. 35th Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, Mar. 2010, Paper OTuD1.
- [32] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.
- [33] H.-C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 90–102, Jan. 2004.
- [34] W. M. Younis, A. H. Sayed, and N. Al-Dhahir, "Efficient adaptive receivers for joint equalization and interference cancellation in multiuser space-time block-coded systems," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2849–2862, Nov. 2003.

**Kohki Shibahara** (Member, IEEE) received the B.S. degree in physics and the M.S. degree in geophysics from Kyoto University in 2008 and 2010, respectively. He joined NTT Network Innovation Laboratories in 2010. He received a Ph.D. in Informatics from Kyoto University in 2017.

His current research interests include spatial division multiplexing transmission systems and advanced multiple-input and multiple-output signal processing. He is a Member of IEICE and the IEEE/Photonics Society. He received the Tingye Li Innovation Prize from OSA in 2016 and the Young Researcher's Award from IEICE in 2017.

**Takayuki Mizuno** (Member, IEEE) received the B.E. degree in applied physics, the M.E. degree in crystalline materials science, and the Dr. Eng. degree in quantum engineering, all from Nagoya University, in 1998, 2000, and 2007, respectively.

In 2000, he joined NTT Photonics Laboratories, NTT Corporation, where he was involved in the research and development of silica planar lightwave circuits. Since 2013, he has been a Senior Research Engineer at NTT Network Innovation Laboratories, NTT Corporation. His current research interests include digital transmission systems and space-division multiplexing transmission technologies for ultra-high capacity photonic networks.

Dr. Mizuno is a member of the OSA, the IEEE Photonics Society, and the IEICE. He has served on or chaired several international conference committees including the Optical Fiber Communication Conference.

**Yutaka Miyamoto** (Member, IEEE) received the B.E. and M.E. degrees in electrical engineering from Waseda University, Tokyo, Japan, in 1986 and 1988, respectively. He received the Dr. Eng. degree in electrical engineering from Tokyo University. He joined the NTT Transmission Systems Laboratories, Yokosuka, Japan, in 1988, where he was engaged in the research and development of high-speed optical communications systems including the 10-Gbit/s first terrestrial optical transmission system (FA-10 G) using EDFA inline repeaters. He then joined NTT Electronics Technology Corporation between 1995 and 1997, where he was engaged in the planning and product development of high-speed optical module at the data rate of 10 Gbps and beyond. Since 1997, he has been with NTT Network Innovation Labs, where he has contributed in the research and development of optical transport technologies based on 40/100/400-Gbit/s channel and beyond. He is a NTT Fellow and the Director of Innovative Photonic Network Research Center, NTT Network Innovation Laboratories, where he has been investigating and promoting the future scalable Optical Transport Network with the Pbit/s-class capacity based on innovative transport technologies such as digital signal processing, space division multiplexing, and cutting-edge integrated devices for photonic pre-processing. He is a member of the IEEE and a Fellow of IEICE.