

JLT-16683-2014-R2
Copyright © 2014 IEEE. This material is posted here with permission of the IEEE. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Polarization-Dependent Loss: New Definition and Measurement Techniques

Reinhold Noé, *Senior Member, IEEE*, Benjamin Koch, David Sandel, *Member, IEEE*, and Vitali Mirvoda

Abstract—We formulate the concatenation properties of polarization-dependent loss (PDL) based on extinction rather than linear units. The advantage of this is that corresponding PDL vectors, defined with length proportional to extinction, can be added with much better accuracy than the traditional linear ones, in particular when PDL is of non-negligible quantity. We also describe either two concatenated PDL elements or a general constant optical element as a combined PDL element and a retarder, thereby obtaining not only input- but also output-referred PDL vectors.

We then propose to model a general optical transmission medium by the concatenation of many differential group delay (DGD) and PDL sections and retarders. An inverse scattering algorithm is provided which allows this physical structure to be obtained from the Jones matrix impulse response. Experimentally, we obtain the latter from the Mueller matrices measured in the optical frequency domain. The finally resulting distributed device structure is displayed in DGD and PDL profiles.

The covariance matrix of the normalized Stokes vectors of scrambled polarizations equals 1/3 times the identity matrix. Based on this, we present yet another PDL measurement technique, the $\sqrt{3}$ scrambling method. It needs no polarimeter and determines low PDL values with better accuracy than the gradient search based extinction method.

Index Terms—Optical fiber communication, optical fiber polarization, polarization-dependent loss, polarization mode dispersion

I. INTRODUCTION

POLARIZATION-dependent loss (PDL) is a fundamental property of optical devices and can not completely be avoided in fiberoptic communication links.

N. Gisin [1] has given the interesting concatenation rules of PDL, based on linear PDL vectors. C. Vinegoni et al. [2] have given separate concatenation rules for the linear PDL units and the transmitted polarization. The same or equivalent PDL vector definitions have been used in [3-6].

A major disadvantage of PDL vectors based on linear units

is that they can be added in the normalized Stokes space only if PDLs are small. In the case of medium and large PDLs the PDL vector addition fails completely when it predicts infinite or complex extinction.

In [5-7], an alternative PDL vector based on extinction units or its derivative with respect to the propagation coordinate are defined. Differential vector equations are also presented. However, no concatenation rule for extinction-based PDL vectors is given. The possibility of expressing extinction-based PDL vector components in dB is pointed out in [8].

In Section II. we use PDL vectors proportional to extinction units and give their concatenation rule. They concatenate much more gracefully than those based on linear units. In particular, the addition of arbitrarily large extinction-based PDL vectors with equal or opposite directions yields the exact result.

We describe two concatenated PDL elements as the combined PDL element plus a retarder, whose parameters are given. In this occasion not only input- but also output-referred PDL vectors are obtained. Furthermore, a general constant optical element is decomposed into the same structure plus a polarization-independent loss element.

The polarization mode dispersion (PMD) or differential group delay (DGD) vectors of concatenated devices can be added in the normalized Stokes space. The sum vector gives the total DGD of the cascade. The vectorial addition can be displayed in a DGD profile [9, 10]. The DGD profile can be determined, together with retarders in between the DGD sections, by measuring the Mueller matrix vs. optical frequency, determining the corresponding Jones matrices, Fourier backtransforming the Jones matrix spectrum into the time domain and inverse scattering [11, 12] of one column vector of the resulting matrix impulse response. Many authors have stated that PMD and PDL are interwoven in typical transmission links. In Section III. we therefore modify the method to an inverse scattering of the full matrix impulse response. That way also the PDL of each DGD section is determined. Just like the DGD vectors, the PDL vectors can be displayed in a PDL profile. Unless PDL is small, it is important to use extinction-based PDL vectors here because their sum is a good approximation of the total PDL vector, a much better one than when the traditional linear PDL vectors are used. Experimental results are given for this distributed PDL and DGD measurement technique.

DGD and PDL profiles give graphical information about the structure of the optical path, not just about its behavior (in

Manuscript received August 12, 2014; revised November 11 and December 22, 2014; accepted December 22, 2014.

R. Noé, B. Koch, D. Sandel and V. Mirvoda are with University of Paderborn, EIM-E, Warburger Str. 100, 33098 Paderborn, Germany (phone: +49 5251 605823, fax: +49 5251 605827, e-mail: noe@upb.de, koch/sandel/mirvoda@ont.upb.de).

R. Noé and B. Koch are also with Novoptel GmbH, EIM-E, Warburger Str. 100, 33098 Paderborn, Germany.

Copyright (c) 2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

frequency or time domain). Only a knowledge of the structure permits, e.g., elimination of high-PDL or high-PMD sections. To measure only PDL at one optical frequency, several input polarizations must be applied to the device under test [13, 14]. In the *Jones, Mueller or Mueller-Jones matrix methods* [13-19] the output polarizations are measured polarimetrically and the input polarizations are known or are likewise measured. PDL is thereby determined fast. Non-polarimetric techniques with mere intensity measurements also exist. Random polarization scrambling with recording of the intensity maximum and minimum is known as the *scanning or all-states method* [19, 20]. For very high PDL extinctions it needs to run for a relatively long time. Very low values are quite susceptible to measurement noise because most intensities are discarded; only two (or a small subset) are being used.

The *maximum-minimum search method* [21, 22] changes polarization transformer settings in a trial-and-error process toward the intensity extrema, which yields PDL. The limitation here is that large steps usually make the user miss the true intensity maximum or minimum whereas small steps make it difficult or impossible to discriminate between two intensities, of which only one is correct. This impedes the search progress.

We go in equal steps to either side of the operation point and thus introduce intensity gradients measured at the operation point into PDL measurement. We call this the *extinction method*. The strategy is well known from optical polarization control [23], where the intensity is maximized or minimized based on such gradients.

In Section IV, we furthermore propose the *sqrt(3) scrambling method*, based on the known correlation properties of scrambled polarizations. We compare the sqrt(3) scrambling against the extinction method experimentally.

Section V. summarizes and concludes this paper.

Regarding notation, we write vectors and matrices in boldface. The tilde (e.g., $\tilde{\mathbf{\Gamma}}$ vs. $\mathbf{\Gamma}$) denotes output- rather than input-referred vectors. The dish-shaped overnotation indicates the 3x3 rotation matrix ($\tilde{\mathbf{W}}$) of a retarder inside the corresponding Mueller matrix (\mathbf{W}).

II. PDL CONCATENATION, DEFINITION, DECOMPOSITION

The traditional [1-6] linear PDL unit Γ_l (Γ in [1]) and associated linear PDL vector $\tilde{\mathbf{\Gamma}}_l$ ($\mathbf{\Gamma}$ in [1]) are

$$\Gamma_l = \frac{T_{\max} - T_{\min}}{T_{\max} + T_{\min}} = |\tilde{\mathbf{\Gamma}}_l|, \quad \tilde{\mathbf{\Gamma}}_l = \Gamma_l \tilde{\mathbf{V}}, \quad (1)$$

where subscript l stands for linear, T_{\max} , T_{\min} are the power transmissions of the strongest and weakest polarization, respectively, and $\tilde{\mathbf{V}}$ is the normalized Stokes vector of the strongest polarization. Gisin's concatenation rule [1] for subsequent devices 1, 2 with linear PDL vectors $\tilde{\mathbf{\Gamma}}_{l,1}$, $\tilde{\mathbf{\Gamma}}_{l,2}$ can be written as

$$\tilde{\mathbf{\Gamma}}_l = \frac{\sqrt{1 - \Gamma_{l,2}^2} \tilde{\mathbf{\Gamma}}_{l,1} + \left(1 + \frac{\tilde{\mathbf{\Gamma}}_{l,2}^T \tilde{\mathbf{\Gamma}}_{l,1}}{\Gamma_{l,2}^2} \cdot \left(1 - \sqrt{1 - \Gamma_{l,2}^2} \right) \right) \tilde{\mathbf{\Gamma}}_{l,2}}{1 + \tilde{\mathbf{\Gamma}}_{l,2}^T \tilde{\mathbf{\Gamma}}_{l,1}}. \quad (2)$$

Here, $\tilde{\mathbf{\Gamma}}_l$ is a linear PDL vector of the cascade and T means transpose.

Vinegoni [2] also gives a concatenation rule for a linear PDL vector $\mathbf{\Gamma}_l = \mathbf{p}_{12}/p_{12}$. Using the auxiliary equations $a_1 = (1/2)(\sqrt{p_1 + |\mathbf{p}_1|} + \sqrt{p_1 - |\mathbf{p}_1|})$, $\mathbf{a}_1 = \mathbf{p}_1/(2a_1)$, it can be calculated from the linear PDL vectors of the individual devices,

$$p_{12} = p_1 p_2 + \mathbf{p}_1^T \mathbf{p}_2 \quad (3)$$

$$\mathbf{p}_{12} = p_1 \mathbf{p}_2 + p_2 \mathbf{p}_1 + 2\mathbf{a}_1 \times (\mathbf{a}_1 \times \mathbf{p}_2)$$

PDL can also be expressed in decibel: PDL [dB] = $10 \log(T_{\max}/T_{\min})$.

We find it useful to work with the extinction unit γ (called a_j in [5, 6]), given by

$$\gamma = \text{PDL}[\text{dB}] \frac{\ln 10}{20} = \frac{1}{2} \ln \frac{T_{\max}}{T_{\min}}, \quad \tanh \gamma = \Gamma_l = |\mathbf{\Gamma}_l| = |\tilde{\mathbf{\Gamma}}_l|. \quad (4)$$

Based on γ we define the *input-referred PDL vector* $\mathbf{\Gamma} = \mathbf{V}\gamma$ (called \mathbf{a}_j in [5, 6]) and the *output-referred PDL vector* $\tilde{\mathbf{\Gamma}} = \tilde{\mathbf{V}}\gamma$. Their derivative (called $\mathbf{a}(z)$) with respect to the propagation coordinate has been used in [5-7]. This is analogous to the definition of input- and output-referred PMD vectors [24, 25]. It holds

$$\mathbf{\Gamma}_l = \mathbf{\Gamma} \frac{\tanh|\mathbf{\Gamma}|}{|\mathbf{\Gamma}|} = \mathbf{V} \tanh \gamma, \quad \tilde{\mathbf{\Gamma}}_l = \tilde{\mathbf{\Gamma}} \frac{\text{arctanh}|\mathbf{\Gamma}_l|}{|\mathbf{\Gamma}_l|} = \tilde{\mathbf{V}} \gamma. \quad (5)$$

Analogous equations allow transforming between $\tilde{\mathbf{\Gamma}}$ and $\tilde{\mathbf{\Gamma}}_l$.

The Jones matrix of a partial polarizer with eigenvalues $\lambda_{1,2} = e^{\pm\gamma/2}$ ($\gamma > 0$) and orthogonal eigenvectors

$$\mathbf{E}_1 = \frac{1}{\sqrt{2(1+V_1)}} \begin{bmatrix} 1+V_1 \\ V_2 - jV_3 \end{bmatrix}, \quad \mathbf{E}_2 = \frac{1}{\sqrt{2(1+V_1)}} \begin{bmatrix} -V_2 - jV_3 \\ 1+V_1 \end{bmatrix}$$

($V_1^2 + V_2^2 + V_3^2 = 1$) is given by

$$\mathbf{J} = \begin{bmatrix} \cosh(\gamma/2) + V_1 \sinh(\gamma/2) & (V_2 + jV_3) \sinh(\gamma/2) \\ (V_2 - jV_3) \sinh(\gamma/2) & \cosh(\gamma/2) - V_1 \sinh(\gamma/2) \end{bmatrix}. \quad (6)$$

\mathbf{J} is Hermitian. The corresponding symmetric Mueller matrix

$$\mathbf{M} = \begin{bmatrix} \cosh \gamma & V_1 \sinh \gamma & V_2 \sinh \gamma & V_3 \sinh \gamma \\ V_1 \sinh \gamma & 1+V_1^2 & V_1 V_2 & V_1 V_3 \\ V_2 \sinh \gamma & V_1 V_2 & 1+V_2^2 & V_2 V_3 \\ V_3 \sinh \gamma & V_1 V_3 & V_2 V_3 & 1+V_3^2 \end{bmatrix} \quad (7)$$

with elements M_{ij} ($i, j = 0 \dots 3$) has eigenvalues $e^{\pm \gamma}$ (= maximum and minimum power transmissions T_{\max} , T_{\min} , here with a geometrical average $T_{\text{ga}} = \sqrt{T_{\max} T_{\min}}$ equal to 1) for Stokes eigenvectors $\mathbf{S} = [1 \pm \mathbf{V}]^T$ ($\mathbf{V} = [V_1 \ V_2 \ V_3]^T$), i.e. for the strongest/weakest polarizations. Input- and output-referred PDL vectors are identical here because \mathbf{M} is

symmetric.

To investigate PDL concatenation, we let the light pass through a first and a second partial polarizer with Mueller matrices \mathbf{M}_i ($i = 1, 2$) of type (7), associated γ_i and $\mathbf{V}_i = [V_{1i} \ V_{2i} \ V_{3i}]^T$. Their product $\mathbf{M}_{21} = \mathbf{M}_2 \mathbf{M}_1$ (8) characterizes the cascaded partial polarizers. It can be decomposed as another matrix product $\mathbf{M}_{21} = \mathbf{W} \mathbf{M}$. Matrix \mathbf{M} of type (7) stands for a combined partial polarizer whose extinction unit and strongest polarization shall be derived in the following. Quantity $\mathbf{W} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{W}} \end{bmatrix}$, with $\tilde{\mathbf{W}}$ being a 3×3 rotation matrix, is the Mueller matrix of a retarder. Its 0th line is $[1 \ 0 \ 0 \ 0]$. So, the 0th line of \mathbf{M}_{21} (or any other matrix which can be factorized as $\mathbf{W} \mathbf{M}$) is identical to that of \mathbf{M} . By equating the 0th lines of (8) and (7) one ends up with

$$\mathbf{M}_{21} = \mathbf{M}_2 \mathbf{M}_1 = \begin{bmatrix} \cosh \gamma_2 \cosh \gamma_1 & \cosh \gamma_2 V_{11} \sinh \gamma_1 & \cosh \gamma_2 V_{21} \sinh \gamma_1 & \cosh \gamma_2 V_{31} \sinh \gamma_1 \\ + V_{12} \sinh \gamma_2 V_{11} \sinh \gamma_1 & + V_{12} \sinh \gamma_2 & + V_{12} \sinh \gamma_2 & + V_{12} \sinh \gamma_2 \\ + V_{22} \sinh \gamma_2 V_{21} \sinh \gamma_1 & \left(1 + V_{11}^2 (\cosh \gamma_1 - 1)\right) & V_{11} V_{21} (\cosh \gamma_1 - 1) & V_{11} V_{31} (\cosh \gamma_1 - 1) \\ + V_{32} \sinh \gamma_2 V_{31} \sinh \gamma_1 & + V_{22} \sinh \gamma_2 & \left(1 + V_{21}^2 (\cosh \gamma_1 - 1)\right) & + V_{22} \sinh \gamma_2 \\ & V_{11} V_{21} (\cosh \gamma_1 - 1) & + V_{32} \sinh \gamma_2 & V_{21} V_{31} (\cosh \gamma_1 - 1) \\ & + V_{32} \sinh \gamma_2 & & + V_{32} \sinh \gamma_2 \\ & V_{11} V_{31} (\cosh \gamma_1 - 1) & V_{21} V_{31} (\cosh \gamma_1 - 1) & \left(1 + V_{31}^2 (\cosh \gamma_1 - 1)\right) \\ \hline V_{12} \sinh \gamma_2 \cosh \gamma_1 & V_{12} \sinh \gamma_2 V_{11} \sinh \gamma_1 & V_{12} \sinh \gamma_2 V_{21} \sinh \gamma_1 & V_{12} \sinh \gamma_2 V_{31} \sinh \gamma_1 \\ + \left(1 + V_{12}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{12}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{12}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{12}^2 (\cosh \gamma_2 - 1)\right) \\ V_{11} \sinh \gamma_1 & \left(1 + V_{11}^2 (\cosh \gamma_1 - 1)\right) & V_{11} V_{21} (\cosh \gamma_1 - 1) & V_{11} V_{31} (\cosh \gamma_1 - 1) \\ + V_{12} V_{22} (\cosh \gamma_2 - 1) & + V_{12} V_{22} (\cosh \gamma_2 - 1) & + V_{12} V_{22} (\cosh \gamma_2 - 1) & + V_{12} V_{22} (\cosh \gamma_2 - 1) \\ V_{21} \sinh \gamma_1 & V_{11} V_{21} (\cosh \gamma_1 - 1) & \left(1 + V_{21}^2 (\cosh \gamma_1 - 1)\right) & V_{21} V_{31} (\cosh \gamma_1 - 1) \\ + V_{12} V_{32} (\cosh \gamma_2 - 1) & + V_{12} V_{32} (\cosh \gamma_2 - 1) & + V_{12} V_{32} (\cosh \gamma_2 - 1) & + V_{12} V_{32} (\cosh \gamma_2 - 1) \\ V_{31} \sinh \gamma_1 & V_{11} V_{31} (\cosh \gamma_1 - 1) & V_{21} V_{31} (\cosh \gamma_1 - 1) & \left(1 + V_{31}^2 (\cosh \gamma_1 - 1)\right) \\ \hline V_{22} \sinh \gamma_2 \cosh \gamma_1 & V_{22} \sinh \gamma_2 V_{11} \sinh \gamma_1 & V_{22} \sinh \gamma_2 V_{21} \sinh \gamma_1 & V_{22} \sinh \gamma_2 V_{31} \sinh \gamma_1 \\ + V_{12} V_{22} (\cosh \gamma_2 - 1) & + V_{12} V_{22} (\cosh \gamma_2 - 1) & + V_{12} V_{22} (\cosh \gamma_2 - 1) & + V_{12} V_{22} (\cosh \gamma_2 - 1) \\ V_{11} \sinh \gamma_1 & \left(1 + V_{11}^2 (\cosh \gamma_1 - 1)\right) & V_{11} V_{21} (\cosh \gamma_1 - 1) & V_{11} V_{31} (\cosh \gamma_1 - 1) \\ + \left(1 + V_{22}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{22}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{22}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{22}^2 (\cosh \gamma_2 - 1)\right) \\ V_{21} \sinh \gamma_1 & V_{11} V_{21} (\cosh \gamma_1 - 1) & \left(1 + V_{21}^2 (\cosh \gamma_1 - 1)\right) & V_{21} V_{31} (\cosh \gamma_1 - 1) \\ + V_{22} V_{32} (\cosh \gamma_2 - 1) & + V_{22} V_{32} (\cosh \gamma_2 - 1) & + V_{22} V_{32} (\cosh \gamma_2 - 1) & + V_{22} V_{32} (\cosh \gamma_2 - 1) \\ V_{31} \sinh \gamma_1 & V_{11} V_{31} (\cosh \gamma_1 - 1) & V_{21} V_{31} (\cosh \gamma_1 - 1) & \left(1 + V_{31}^2 (\cosh \gamma_1 - 1)\right) \\ \hline V_{32} \sinh \gamma_2 \cosh \gamma_1 & V_{32} \sinh \gamma_2 V_{11} \sinh \gamma_1 & V_{32} \sinh \gamma_2 V_{21} \sinh \gamma_1 & V_{32} \sinh \gamma_2 V_{31} \sinh \gamma_1 \\ + V_{12} V_{32} (\cosh \gamma_2 - 1) & + V_{12} V_{32} (\cosh \gamma_2 - 1) & + V_{12} V_{32} (\cosh \gamma_2 - 1) & + V_{12} V_{32} (\cosh \gamma_2 - 1) \\ V_{11} \sinh \gamma_1 & \left(1 + V_{11}^2 (\cosh \gamma_1 - 1)\right) & V_{11} V_{21} (\cosh \gamma_1 - 1) & V_{11} V_{31} (\cosh \gamma_1 - 1) \\ + V_{22} V_{32} (\cosh \gamma_2 - 1) & + V_{22} V_{32} (\cosh \gamma_2 - 1) & + V_{22} V_{32} (\cosh \gamma_2 - 1) & + V_{22} V_{32} (\cosh \gamma_2 - 1) \\ V_{21} \sinh \gamma_1 & V_{11} V_{21} (\cosh \gamma_1 - 1) & \left(1 + V_{21}^2 (\cosh \gamma_1 - 1)\right) & V_{21} V_{31} (\cosh \gamma_1 - 1) \\ + \left(1 + V_{32}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{32}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{32}^2 (\cosh \gamma_2 - 1)\right) & + \left(1 + V_{32}^2 (\cosh \gamma_2 - 1)\right) \\ V_{31} \sinh \gamma_1 & V_{11} V_{31} (\cosh \gamma_1 - 1) & V_{21} V_{31} (\cosh \gamma_1 - 1) & \left(1 + V_{31}^2 (\cosh \gamma_1 - 1)\right) \end{bmatrix} \quad (8)$$

$$\begin{aligned} \cosh \gamma &= M_{00} = \cosh \gamma_2 \cosh \gamma_1 + \mathbf{V}_2^T \mathbf{V}_1 \sinh \gamma_2 \sinh \gamma_1 \\ \mathbf{V} \sinh \gamma &= [M_{01} \ M_{02} \ M_{03}]^T \\ &= \mathbf{V}_2 \sinh \gamma_2 \\ &+ \mathbf{V}_1 \left(\cosh \gamma_2 \sinh \gamma_1 + \mathbf{V}_2^T \mathbf{V}_1 \sinh \gamma_2 (\cosh \gamma_1 - 1) \right) \\ \mathbf{V} \tanh \gamma &= \frac{\mathbf{V}_2 \sinh \gamma_2 + \mathbf{V}_1 \left(\cosh \gamma_2 \sinh \gamma_1 + \mathbf{V}_2^T \mathbf{V}_1 \sinh \gamma_2 (\cosh \gamma_1 - 1) \right)}{\cosh \gamma_2 \cosh \gamma_1 + \mathbf{V}_2^T \mathbf{V}_1 \sinh \gamma_2 \sinh \gamma_1} \end{aligned} \quad (9)$$

and subsequently $\mathbf{V} = (\mathbf{V} \tanh \gamma) / |\mathbf{V} \tanh \gamma|$ and $\gamma = \operatorname{atanh} |\mathbf{V} \tanh \gamma|$.

Without loss of generality we can assume $\gamma_1 > 0$, $\gamma_2 > 0$, $\gamma > 0$ because negative signs could be taken into account by changing the polarity of \mathbf{V}_1 , \mathbf{V}_2 , \mathbf{V} . Using $\sinh \gamma_2 > 0$, $\cosh \gamma_2 > \sinh \gamma_2$, $\sinh \gamma_1 > \cosh \gamma_1 - 1$ we see that the factors by which \mathbf{V}_1 , \mathbf{V}_2 are multiplied to form $\mathbf{V} \sinh \gamma$ (9) (or \mathbf{V}) as their linear combination are both positive. So, \mathbf{V} lies on a great circle of the Poincaré sphere in between \mathbf{V}_1 , \mathbf{V}_2 , i.e. on the shorter of the two possible circle segments between \mathbf{V}_1 , \mathbf{V}_2 .

When the input-referred PDL vector or corresponding normalized vector is fed into a device, then the output-referred one is generated,

$$\begin{bmatrix} 1 \\ \pm \tilde{\mathbf{V}} \end{bmatrix} e^{\pm \gamma} = \mathbf{W} \mathbf{M} \begin{bmatrix} 1 \\ \pm \mathbf{V} \end{bmatrix} = \mathbf{W} \begin{bmatrix} 1 \\ \pm \mathbf{V} \end{bmatrix} e^{\pm \gamma}. \quad (11)$$

As a consequence it holds $\tilde{\mathbf{V}} = \tilde{\mathbf{W}} \mathbf{V}$, $\tilde{\Gamma} = \tilde{\mathbf{W}} \Gamma$. This is obviously valid for any retarder matrix \mathbf{W} , not just for the one we need in the decomposition of \mathbf{M}_{21} .

We can also factorize $\mathbf{M}_{21} = \tilde{\mathbf{M}} \mathbf{W}$ with $\tilde{\mathbf{M}} = \mathbf{W} \mathbf{M} \mathbf{W}^T$. The elements of $\tilde{\mathbf{M}}$ are \tilde{M}_{ij} . Given that the 0th column of \mathbf{W} is $[1 \ 0 \ 0 \ 0]^T$, the 0th column of \mathbf{M}_{21} (or any other matrix which can be factorized as $\tilde{\mathbf{M}} \mathbf{W}$) equals that of $\tilde{\mathbf{M}}$. One obtains

$$\begin{aligned} \tilde{\mathbf{V}} \sinh \gamma &= [\tilde{M}_{10} \ \tilde{M}_{20} \ \tilde{M}_{30}]^T \\ &= \mathbf{V}_1 \sinh \gamma_1 + \mathbf{V}_2 \left(\cosh \gamma_1 \sinh \gamma_2 \right. \\ &\quad \left. + \mathbf{V}_2^T \mathbf{V}_1 \sinh \gamma_1 (\cosh \gamma_2 - 1) \right). \end{aligned} \quad (12)$$

Since \mathbf{M}_{21} is not symmetric, \mathbf{V} and $\tilde{\mathbf{V}}$ do not coincide. In the same manner as above, we can show that $\tilde{\mathbf{V}}$ lies on the shorter great circle segment between \mathbf{V}_1 , \mathbf{V}_2 .

We can expand vector $\mathbf{V} \times \tilde{\mathbf{V}}$ and will find that it is parallel, with the same sign, to vector $\mathbf{V}_1 \times \mathbf{V}_2$. For $|\mathbf{V}_2^T \mathbf{V}_1| < 1$

we can furthermore show $\mathbf{V}_1^T (\mathbf{V} - \tilde{\mathbf{V}}) \sinh \gamma > 0$, $\mathbf{V}_2^T (\tilde{\mathbf{V}} - \mathbf{V}) \sinh \gamma > 0$. So, \mathbf{V} , $\tilde{\mathbf{V}}$ lie, in this order, between \mathbf{V}_1 , \mathbf{V}_2 (order \mathbf{V}_1 , \mathbf{V} , $\tilde{\mathbf{V}}$, \mathbf{V}_2). The eigenmodes of the rotation submatrix $\tilde{\mathbf{W}}$ lie in a first plane defined by $\mathbf{V} \times \tilde{\mathbf{V}}$, or the parallel vector $\mathbf{V}_1 \times \mathbf{V}_2$, and $\mathbf{V} + \tilde{\mathbf{V}}$.

For symmetric \mathbf{M} (7) with arbitrary input polarization, given by the normalized Stokes vector \mathbf{S} , one obtains

$$\mathbf{M} \begin{bmatrix} 1 \\ \mathbf{S} \end{bmatrix} = \begin{bmatrix} \cosh \gamma + \mathbf{V}^T \mathbf{S} \sinh \gamma \\ \mathbf{V} (\sinh \gamma + \mathbf{V}^T \mathbf{S} (\cosh \gamma - 1)) + \mathbf{S} \end{bmatrix}. \quad (13)$$

The output polarization (or, more precisely, its normalized Stokes vector) is a linear combination of \mathbf{S} and \mathbf{V} . The same way we calculate

$$\begin{aligned} \mathbf{M}_{21} \begin{bmatrix} 1 \\ \pm \mathbf{V}_1 \end{bmatrix} &= \mathbf{M}_2 \mathbf{M}_1 \begin{bmatrix} 1 \\ \pm \mathbf{V}_1 \end{bmatrix} = \mathbf{M}_2 \begin{bmatrix} 1 \\ \pm \mathbf{V}_1 \end{bmatrix} e^{\pm \gamma_1} \\ &= \begin{bmatrix} \cosh \gamma_2 \pm \mathbf{V}_2^T \mathbf{V}_1 \sinh \gamma_2 \\ \mathbf{V}_2 (\sinh \gamma_2 \pm \mathbf{V}_2^T \mathbf{V}_1 (\cosh \gamma_2 - 1)) \pm \mathbf{V}_1 \end{bmatrix} e^{\pm \gamma_1}. \end{aligned} \quad (14)$$

We define $\mathbf{M} \begin{bmatrix} 1 \\ \mathbf{V}_1 \end{bmatrix} = c \begin{bmatrix} 1 \\ \mathbf{C} \end{bmatrix}$, $\mathbf{M}_{21} \begin{bmatrix} 1 \\ \mathbf{V}_1 \end{bmatrix} = d \begin{bmatrix} 1 \\ \mathbf{D} \end{bmatrix}$ where c , d are constants. \mathbf{V}_1 , \mathbf{V}_2 , \mathbf{V} , $\tilde{\mathbf{V}}$, \mathbf{C} , \mathbf{D} all lie in the same plane. It holds $\tilde{\mathbf{V}} = \tilde{\mathbf{W}} \mathbf{V}$ and $\mathbf{D} = \tilde{\mathbf{W}} \mathbf{C}$. Hence the eigenmodes of $\tilde{\mathbf{W}}$ must also lie in a second plane, defined by $\mathbf{C} \times \mathbf{D}$, or the parallel vector $\mathbf{V}_1 \times \mathbf{V}_2$, and $\mathbf{C} + \mathbf{D}$.

It is straightforward to show $\mathbf{V} \times \mathbf{C} = \tilde{\mathbf{V}} \times \mathbf{D}$. This means the two planes have only one common line, $\mathbf{V}_1 \times \mathbf{V}_2$. The eigenmodes of $\tilde{\mathbf{W}}$ are parallel to $\mathbf{V}_1 \times \mathbf{V}_2$!

In contrast, if the eigenmodes of $\tilde{\mathbf{W}}$ were not parallel to $\mathbf{V}_1 \times \mathbf{V}_2$, they would need to be parallel to $\mathbf{V} + \tilde{\mathbf{V}}$ and $\mathbf{C} + \mathbf{D}$. $\mathbf{V} \times \mathbf{C}$ would need to be equal to $-\tilde{\mathbf{V}} \times \mathbf{D}$, but this is not the case!

So, to fulfill $\tilde{\mathbf{V}} = \tilde{\mathbf{W}} \mathbf{V}$, the rotation matrix

$$\tilde{\mathbf{W}} = \begin{bmatrix} B_1^2 + (B_2^2 + B_3^2) & B_1 B_2 (1 - \cos \xi) & B_1 B_3 (1 - \cos \xi) \\ \cdot \cos \xi & -B_3 \sin \xi & +B_2 \sin \xi \\ \hline B_1 B_2 (1 - \cos \xi) & B_2^2 + (B_1^2 + B_3^2) & B_2 B_3 (1 - \cos \xi) \\ +B_3 \sin \xi & \cdot \cos \xi & -B_1 \sin \xi \\ \hline B_1 B_3 (1 - \cos \xi) & B_2 B_3 (1 - \cos \xi) & B_3^2 + (B_1^2 + B_2^2) \\ -B_2 \sin \xi & +B_1 \sin \xi & \cdot \cos \xi \end{bmatrix} \quad (15)$$

turns normalized Stokes vectors about the eigenmode axis $\mathbf{B} = [B_1 \ B_2 \ B_3]^T = (\mathbf{v} \times \tilde{\mathbf{v}}) / |\mathbf{v} \times \tilde{\mathbf{v}}| = (\mathbf{V}_1 \times \mathbf{V}_2) / |\mathbf{V}_1 \times \mathbf{V}_2|$ with a retardation angle ξ that is given by $\cos \xi = \tilde{\mathbf{V}}^T \mathbf{V}$, $\sin \xi = |\mathbf{V} \times \tilde{\mathbf{V}}|$, $\xi = \arctan \left(|\mathbf{V} \times \tilde{\mathbf{V}}| / (\tilde{\mathbf{V}}^T \mathbf{V}) \right)$.

An alternative, lengthy proof of the PDL concatenation rule can be led by inserting the given quantities into

$$\mathbf{M}_2\mathbf{M}_1 = \mathbf{W}\mathbf{M}.$$

Let us look at another configuration. If cascaded partial polarizers are separated by a retarder \mathbf{R} , e.g. with the total matrix product $\mathbf{N} = \mathbf{M}_2\mathbf{R}\mathbf{M}_1$, then the PDL vectors and the symmetric partial polarizer matrices must all be transformed to the same location, input ($\mathbf{N} = \mathbf{R}(\mathbf{R}^T\mathbf{M}_2\mathbf{R})\mathbf{M}_1$) or output ($\mathbf{N} = \mathbf{M}_2(\mathbf{R}\mathbf{M}_1\mathbf{R}^T)\mathbf{R}$).

For small PDL, matters are considerably simpler. In (9) and elsewhere one may set $\cosh \gamma_i \approx 1$, $\sinh \gamma_i \approx \gamma_i \approx \tanh \gamma_i$, thus obtaining

$$\begin{aligned} \mathbf{V}\gamma &= \mathbf{V}_1\gamma_1 + \mathbf{V}_2\gamma_2 & \mathbf{\Gamma} &= \mathbf{\Gamma}_1 + \mathbf{\Gamma}_2 \\ \tilde{\mathbf{V}}\gamma &= \tilde{\mathbf{V}}_1\gamma_1 + \tilde{\mathbf{V}}_2\gamma_2 & \tilde{\mathbf{\Gamma}} &= \tilde{\mathbf{\Gamma}}_1 + \tilde{\mathbf{\Gamma}}_2 \end{aligned} \quad (|\gamma_i| \ll 1). \quad (16)$$

All the foregoing can be executed several times sequentially if more than two partial polarizers are cascaded. As exemplified above, all PDL vectors must be referred to the same location, i.e. they must belong to adjacent symmetric Mueller matrices.

Transformation between extinction-based and linear PDL vectors is easy (5). This way one can make the earlier-given concatenation rules (2), (3) for linear PDL vectors usable for extinction-based ones. Remember that input- and output-referred PDL vectors are identical only for symmetric Mueller matrices. We found that Gisin [1] returns the output-referred linear PDL vector $\tilde{\mathbf{\Gamma}}_l$ (2) whereas Vinegoni [2] delivers the ingredients of the input-referred linear PDL vector $\mathbf{\Gamma}_l = \mathbf{p}_{12}/p_{12}$ (3).

When the different definitions are considered, all concatenation rules (Gisin, Vinegoni, ours) deliver the same results. For small PDL, (2), (3) can be replaced by a simple addition $\tilde{\mathbf{\Gamma}}_l = \tilde{\mathbf{\Gamma}}_{l,1} + \tilde{\mathbf{\Gamma}}_{l,2}$ or $\mathbf{\Gamma}_l = \mathbf{\Gamma}_{l,1} + \mathbf{\Gamma}_{l,2}$. However, due to $|\mathbf{\Gamma}_l| \leq 1$ (which follows from $T_{\min} \geq 0$) this gets catastrophically wrong for $|\mathbf{\Gamma}_{l,1} + \mathbf{\Gamma}_{l,2}| > 1$, and even for moderate PDL the error can be significant.

Regarding the direct addition of PDL vectors, their definition based on the extinction parameter γ is superior to the linear one if PDL is not small. The striking difference is illustrated in two examples:

1) We take two equally oriented polarizers with extinction units γ_1, γ_2 . Adding our PDL vectors $\mathbf{\Gamma}_1, \mathbf{\Gamma}_2$ (which is, in this case $\mathbf{V}_1 = \mathbf{V}_2$, the direct result of (9), (10)!) gives the correct $\mathbf{\Gamma} = \mathbf{\Gamma}_1 + \mathbf{\Gamma}_2$ and the correct extinction unit $\gamma = \gamma_1 + \gamma_2$. Everyone knows indeed that the extinctions of cascaded polarizers can be added! However, the addition of the corresponding linear PDL vectors rather than the application of the correct concatenation rule gives, after back-conversion, an extinction unit $\text{arctanh}(\tanh \gamma_1 + \tanh \gamma_2)$, which differs from the correct value $\gamma_1 + \gamma_2$. For individual extinctions as small as $\gamma_1 = \gamma_2 = 0.213$ (corresponding to 1.85 dB) the relative extinction error which results from

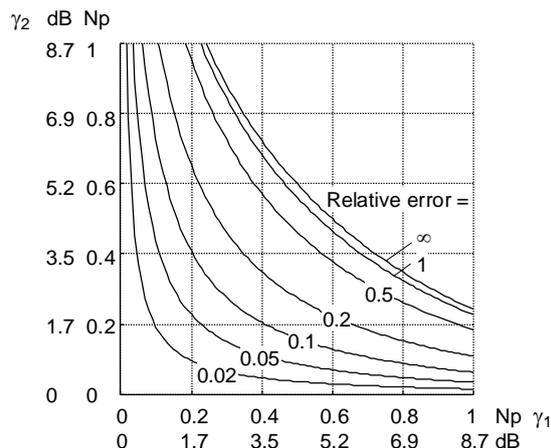


Fig. 1. Contour plot of relative error (1 means 100%) resulting from linear PDL vector addition $\text{arctanh}(\tanh \gamma_1 + \tanh \gamma_2)$ vs. extinction-based PDL vector addition $\gamma_1 + \gamma_2$ for equally oriented partial polarizers. Extinctions γ_1, γ_2 are given dimensionless in Neper (Np) and in dB.

addition of linear PDL vectors reaches 5%, and it becomes infinite for modest $\gamma_1 = \gamma_2 = 0.55$ (corresponding to 4.8 dB). More generally, Fig. 1 shows the contours of the relative error $(\text{arctanh}(\tanh \gamma_1 + \tanh \gamma_2))/(\gamma_1 + \gamma_2) - 1$ for arbitrary combinations of γ_1, γ_2 .

2) Consider two cascaded PDL elements of type (7) with equal extinctions of 7.7 dB, corresponding $\gamma_1 = \gamma_2 \approx 0.886$, and $0^\circ/90^\circ$ ($\mathbf{V}_1 = [1 \ 0 \ 0]^T$) and $\pm 45^\circ$ ($\mathbf{V}_2 = [0 \ 1 \ 0]^T$) eigenmodes (= strongest/weakest polarizations). When adding our PDL vectors $\mathbf{\Gamma}_i = \mathbf{V}_i\gamma_i$ according to (16) we get $\gamma = |\mathbf{\Gamma}| = \sqrt{2}\gamma_{1,2} \approx 1.253$ and a total extinction of 10.9 dB. This is only 5% off the correct value 11.5 dB calculated with (9), (10). In contrast, the addition of the corresponding traditionally defined PDL vectors [1, 2] $\mathbf{\Gamma}_{l,i} = \mathbf{V}_i \tanh \gamma_i$ with lengths $\tanh \gamma_i \approx 0.709$ predicts, after back-conversion, a complex (or, if one limits the argument of the arctanh , an infinite) extinction!

Let us repeat that both PDL vector definitions are correct and are flanked by correct concatenation rules. The extinction-based PDL vectors turn out to be advantageous (i.e. much more accurate) when one adds PDL vectors instead of properly concatenating them. This simple PDL vector addition will indeed be needed in the PDL profile plotting of Section III. For accuracy reasons we use extinction-based PDL vectors there.

Returning to matrix decomposition, the 7 elements contained in the 0th line and 0th column of \mathbf{M}_{21} contain just 5 degrees-of-freedom (DOF) for \mathbf{V} (2 DOF), $\tilde{\mathbf{V}}$ (2 DOF), γ (1 DOF). From these, eigenmodes and retardation of $\tilde{\mathbf{W}}$ are derived. We therefore describe how a general constant element, whose Mueller matrix \mathbf{P} has 7 DOF (= 8 DOF of a Jones matrix minus 1 DOF representing the common phase), can be expressed by a retarder (3 DOF), a PDL element of

type (7) (3 DOF) and a polarization-independent power transmission (1 DOF). For Jones matrices, this possibility was already pointed out in [6].

For Mueller matrix \mathbf{P} , this starts with the singular value decomposition

$$\mathbf{P} = \mathbf{U}\mathbf{T}\mathbf{V}^T \quad (17)$$

which yields orthogonal matrices \mathbf{U} , \mathbf{V} and a diagonal singular value matrix \mathbf{T} , containing transmissions. The elements of \mathbf{T} are made non-negative and put in decreasing order. This is achieved by sign changes and reordering of the column vectors of \mathbf{U} , \mathbf{V} . Matrix \mathbf{T} now contains the four non-negative singular values in decreasing order, namely T_{\max} , T_{ga} , T_{ga} , T_{\min} . Two of these are identical and are the geometric average of the others, $T_{\max} \geq T_{\text{ga}} = \sqrt{T_{\max}T_{\min}} \geq T_{\min}$.

The singular values T_{\max} , T_{\min} are indeed the power transmissions of the strongest and weakest polarization, respectively. We may choose the latter to be $0^\circ/90^\circ$ linear (along the S_1 axis). T_{ga} is also the transmission of two other Stokes parameters, here S_2 and $\pm S_3$, which define a cartesian coordinate system for the Poincaré sphere together with the strongest/weakest polarization axis S_1 . So the diagonal matrix \mathbf{T} does not describe the transmission of a Stokes vector. Rather it describes the transmissions of the components $|E_x|^2$ (= power in x-component), S_2 , $\pm S_3$, and $|E_y|^2$ (= power in y-component). We can express the Stokes vector through

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 0 & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \pm 1 & 0 \end{bmatrix}}_{\mathbf{L}} \begin{bmatrix} \sqrt{2}|E_x|^2 \\ S_2 \\ \pm S_3 \\ \sqrt{2}|E_y|^2 \end{bmatrix} \quad (\mathbf{L}^{-1} = \mathbf{L}^T) \quad (18)$$

with orthogonal \mathbf{L} . In the equation

$$\mathbf{P} = \underbrace{\mathbf{U}\mathbf{L}^T}_{\mathbf{R}_2} \underbrace{\mathbf{L}\mathbf{T}\mathbf{L}^T}_{T_{\text{ga}}\mathbf{M}_0} \underbrace{\mathbf{L}\mathbf{V}^T}_{\mathbf{R}_1}, \quad (19)$$

obviously the symmetric matrix $T_{\text{ga}}\mathbf{M}_0$ represents a partial $0^\circ/90^\circ$ polarizer with geometric average power transmission T_{ga} and a \mathbf{M}_0 conforming to (7) with $T_{\max}/T_{\min} = e^{2\gamma}$,

$\mathbf{V} = [1 \ 0 \ 0]^T$. The two orthogonal matrices \mathbf{R}_1 , \mathbf{R}_2 must represent retarders. For the latter to be true, $\det \mathbf{R}_1 = \det \mathbf{R}_2 = 1$ must hold. Even though the sign of $\det \mathbf{U} = \det \mathbf{V}^T = \pm 1$ is usually undetermined, one can meet this condition by choosing either upper or lower sign in (18), which results in $\det \mathbf{L} = \pm 1$. The first row of the 3×3 rotation submatrix of type (15) inside \mathbf{R}_1 is the normalized input-referred PDL vector \mathbf{V} , and the first column of the rotation

submatrix inside \mathbf{R}_2 is the normalized output-referred PDL vector $\tilde{\mathbf{V}}$.

Now one may write

$$\mathbf{P} = T_{\text{ga}} \underbrace{\mathbf{R}_2 \mathbf{R}_1}_{\text{retarder}} \underbrace{\mathbf{R}_1^T \mathbf{M}_0 \mathbf{R}_1}_{\text{partial pol.}} = T_{\text{ga}} \underbrace{\mathbf{R}_2 \mathbf{M}_0 \mathbf{R}_2^T}_{\text{partial pol.}} \underbrace{\mathbf{R}_2 \mathbf{R}_1}_{\text{retarder}} \quad (20)$$

which is the desired decomposition. The retarders are lossless, the partial polarizer matrices symmetric of type (7). We have numerically confirmed this decomposition procedure.

III. DGD AND PDL PROFILES DETERMINED BY INVERSE SCATTERING

Most authors express the PMD vector by a Taylor series [26]. This gives unphysical results (infinite PMD!) far off the carrier frequency. That effect is not surprising, because the PMD vector varies over frequency in a quasi-periodic manner, and a Taylor series is usually a bad approximation for periodic phenomena.

In contrast, several DGD sections, which can be plotted in a DGD profile, are a physical and very effective description of a general PMD medium [9, 10]. The DGD profile graphically adds individual PMD vectors so that the sum represents the overall 1st-order PMD.

The individual DGD sections are separated by retarders [27]. In later work this was called a hinge model [28]. It is important to note that only one of the retarders needs to be a general elliptical retarder [27]. Assuming DGD sections with S_1 ($= 0^\circ/90^\circ$) eigenmodes the other retarders can be Soleil-Babinet analogs (SBAs) [27], having eigenmodes which can be oriented freely on the $S_2 - S_3$ great circle. They can also be called in-phase and quadrature mode converters. To get the rotation matrix of an SBA with retardation φ_i and orientation

ψ_i one sets $\xi = \varphi_i$ and $\mathbf{B} = [0 \ \cos \psi_i \ \sin \psi_i]^T$ in (15). The corresponding Jones matrix is given in the following eqn. (23). An inverse scattering process to determine these components from the optical vector impulse response has been outlined in [11, 12] and demonstrated in [9].

Extending this model, we describe the transmission path in $n+1$ steps by

- (0) Element with constant Jones matrix,
- (1) DGD section followed by SBA and PDL,
- ...
- (n) DGD section followed by SBA and PDL.

The i -th PDL element and the $(i+1)$ -th DGD section have the same orientation with $0^\circ/90^\circ$ eigenmodes and could, as a consequence, be exchanged. Together they form a partial polarizer and phase shifter with extinction parameter γ_i and retardation $\delta = \omega\tau$. Moreover, a phase shifter, which is contained in the element of step (0), can be exchanged with a DGD section and also with an SBA if the orientation of the latter is changed by the phase shifter's retardation [27]. So, the SBAs in the steps (1) to (n) could just as well be elliptical retarders.

Let \mathbf{k}_i be an intermediate result after DGD, but before SBA, \mathbf{l}_i be an intermediate result after SBA but before PDL, and \mathbf{h}_i be the matrix impulse response to the unity excitation matrix $\mathbf{1} \cdot \delta(t)$, each of these in or after the i -th step. Using \mathbf{g} for \mathbf{h} , \mathbf{k} , \mathbf{l} we write

$$\mathbf{g}_i = \begin{bmatrix} g_{11,0,i} & g_{12,0,i} \\ g_{21,0,i} & g_{22,0,i} \end{bmatrix} \delta(t) + \begin{bmatrix} g_{11,1,i} & g_{12,1,i} \\ g_{21,1,i} & g_{22,1,i} \end{bmatrix} \delta(t - \tau) + \dots + \begin{bmatrix} g_{11,i,i} & g_{12,i,i} \\ g_{21,i,i} & g_{22,i,i} \end{bmatrix} \delta(t - i\tau) \quad (21)$$

The 1st and 2nd columns of \mathbf{g} contain the vector responses to horizontal and vertical polarization impulses. The passing of the signal through the elements $i > 0$ is described by

$$\mathbf{k}_i = \underbrace{\begin{bmatrix} \delta(t) & 0 \\ 0 & \delta(t - \tau) \end{bmatrix}}_{\text{DGD}(\tau)} * \mathbf{h}_{i-1}, \quad (22)$$

$$\mathbf{l}_i = \underbrace{\begin{bmatrix} \cos(\varphi_i/2) & je^{j\psi_i} \sin(\varphi_i/2) \\ je^{-j\psi_i} \sin(\varphi_i/2) & \cos(\varphi_i/2) \end{bmatrix}}_{\text{SBA}(\varphi_i, \psi_i)} \cdot \mathbf{k}_i, \quad (23)$$

$$\mathbf{h}_i = \underbrace{\begin{bmatrix} e^{\gamma_i/2} & 0 \\ 0 & e^{-\gamma_i/2} \end{bmatrix}}_{\text{PDL}(\gamma_i)} \cdot \mathbf{l}_i. \quad (24)$$

Starting from

$$\mathbf{h}_0 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \delta(t), \quad (25)$$

this yields successively

$$\mathbf{k}_i = \begin{bmatrix} h_{11,0,i-1} & h_{12,0,i-1} \\ 0 & 0 \end{bmatrix} \delta(t) + \dots + \begin{bmatrix} 0 & 0 \\ h_{21,i-1,i-1} & h_{22,i-1,i-1} \end{bmatrix} \delta(t - i\tau), \quad (26)$$

\mathbf{l}_i (27) and \mathbf{h}_i (28).

For inverse scattering we let i run from n down to 1. Given that some elements can vanish in practice it is useful to weight them quadratically. The PDL parameter is given by

$$\frac{|h_{11,0,i}|^2 + |h_{12,0,i}|^2}{|h_{21,0,i}|^2 + |h_{22,0,i}|^2} \cdot \frac{|h_{21,i,i}|^2 + |h_{22,i,i}|^2}{|h_{11,i,i}|^2 + |h_{12,i,i}|^2} = e^{4\gamma_i} \quad (29)$$

Knowing γ_i we can calculate

$$\mathbf{l}_i = \text{PDL}(-\gamma_i) \cdot \mathbf{h}_i. \quad (30)$$

Furthermore we see

$$-\frac{l_{21,0,i}^*}{l_{11,0,i}^*} = -\frac{l_{22,0,i}^*}{l_{12,0,i}^*} = \frac{l_{11,i,i}}{l_{21,i,i}} = \frac{l_{12,i,i}}{l_{22,i,i}} = je^{j\psi_i} \tan(\varphi_i/2). \quad (31)$$

For weighting in (31) we consider $\sin(\varphi_i/2) = l_{11,i,i} / \sqrt{|l_{11,i,i}|^2 + |l_{21,i,i}|^2}$ etc. and evaluate $\sin(\varphi_i/2)$, $\cos(\varphi_i/2)$ by (32). This yields the SBA retardation φ_i . With weighting, the SBA orientation is computed as

$$\psi_i = \arg \left(\frac{-\left(l_{21,0,i} \cdot l_{11,0,i}^*\right) - \left(l_{22,0,i} \cdot l_{12,0,i}^*\right)}{+l_{11,i,i} \cdot l_{21,i,i}^* + l_{12,i,i} \cdot l_{22,i,i}^*} \right) - \pi/2. \quad (33)$$

Using the SBA parameters one obtains

$$\mathbf{l}_i = \begin{bmatrix} \cos(\varphi_i/2)h_{11,0,i-1} & \cos(\varphi_i/2)h_{12,0,i-1} \\ je^{-j\psi_i} \sin(\varphi_i/2)h_{11,0,i-1} & je^{-j\psi_i} \sin(\varphi_i/2)h_{12,0,i-1} \end{bmatrix} \delta(t) + \dots + \begin{bmatrix} je^{j\psi_i} \sin(\varphi_i/2)h_{21,i-1,i-1} & je^{j\psi_i} \sin(\varphi_i/2)h_{22,i-1,i-1} \\ \cos(\varphi_i/2)h_{21,i-1,i-1} & \cos(\varphi_i/2)h_{22,i-1,i-1} \end{bmatrix} \delta(t - i\tau) \quad (27)$$

$$\mathbf{h}_i = \begin{bmatrix} e^{\gamma_i/2} \cos(\varphi_i/2)h_{11,0,i-1} & e^{\gamma_i/2} \cos(\varphi_i/2)h_{12,0,i-1} \\ je^{-\gamma_i/2 - j\psi_i} \sin(\varphi_i/2)h_{11,0,i-1} & je^{-\gamma_i/2 - j\psi_i} \sin(\varphi_i/2)h_{12,0,i-1} \end{bmatrix} \delta(t) + \dots + \begin{bmatrix} je^{\gamma_i/2 + j\psi_i} \sin(\varphi_i/2)h_{21,i-1,i-1} & je^{\gamma_i/2 + j\psi_i} \sin(\varphi_i/2)h_{22,i-1,i-1} \\ e^{-\gamma_i/2} \cos(\varphi_i/2)h_{21,i-1,i-1} & e^{-\gamma_i/2} \cos(\varphi_i/2)h_{22,i-1,i-1} \end{bmatrix} \delta(t - i\tau) \quad (28)$$

$$\sin(\varphi_i/2) = \frac{|l_{21,0,i}| \sqrt{|l_{21,0,i}|^2 + |l_{11,0,i}|^2} + |l_{22,0,i}| \sqrt{|l_{22,0,i}|^2 + |l_{12,0,i}|^2} + |l_{11,i,i}| \sqrt{|l_{11,i,i}|^2 + |l_{21,i,i}|^2} + |l_{12,i,i}| \sqrt{|l_{12,i,i}|^2 + |l_{22,i,i}|^2}}{|l_{21,0,i}|^2 + |l_{11,0,i}|^2 + |l_{22,0,i}|^2 + |l_{12,0,i}|^2 + |l_{11,i,i}|^2 + |l_{21,i,i}|^2 + |l_{12,i,i}|^2 + |l_{22,i,i}|^2}} \quad (32)$$

$$\cos(\varphi_i/2) = \frac{|l_{11,0,i}| \sqrt{|l_{21,0,i}|^2 + |l_{11,0,i}|^2} + |l_{12,0,i}| \sqrt{|l_{22,0,i}|^2 + |l_{12,0,i}|^2} + |l_{21,i,i}| \sqrt{|l_{11,i,i}|^2 + |l_{21,i,i}|^2} + |l_{22,i,i}| \sqrt{|l_{12,i,i}|^2 + |l_{22,i,i}|^2}}{|l_{21,0,i}|^2 + |l_{11,0,i}|^2 + |l_{22,0,i}|^2 + |l_{12,0,i}|^2 + |l_{11,i,i}|^2 + |l_{21,i,i}|^2 + |l_{12,i,i}|^2 + |l_{22,i,i}|^2}}$$

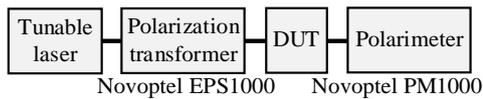


Fig. 2. Setup for DGD and PDL profile measurement

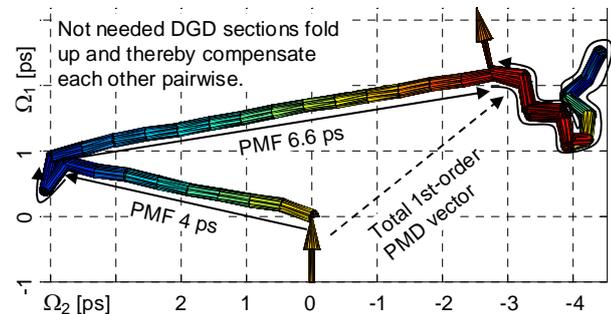


Fig. 3. DGD profile of two randomly concatenated pieces of PMF with 4 and 6.6 ps of DGD. Not needed DGD sections, which fold up and thereby compensate each other pairwise, are marked by curved arrows.

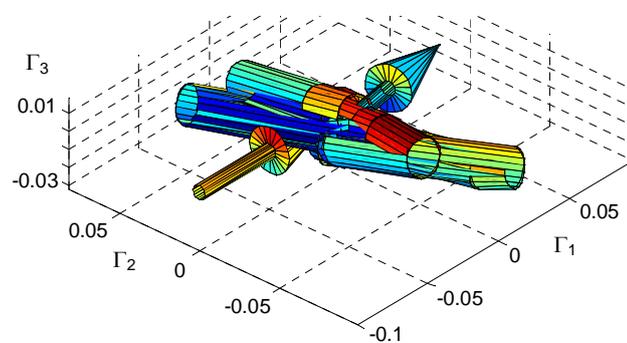


Fig. 4. PDL profile, containing extinction-based PDL vectors, of two randomly concatenated pieces of PMF with 4 and 6.6 ps of DGD

$$\mathbf{h}_{i-1} = \text{DGD}(-\tau) * \text{SBA}(-\varphi_i, \psi_i) \cdot \mathbf{l}_i. \quad (34)$$

The final step provides us with \mathbf{h}_0 . For analysis of the 0th optical element with constant Jones matrix $\int \mathbf{h}_0 dt$ one can extract its eigenvalues and not necessarily orthogonal eigenmodes. One can also determine the corresponding Mueller matrix and factorize it according to (20).

Experimentally it makes sense to limit $|\gamma_i|$ or to set $\gamma_i = 0$ when the elements of the outer matrices in \mathbf{h}_i with indexes $\dots, 0, i$ and \dots, i, i , or their determinant, are very small in magnitude and therefore contain significant relative errors. Also, if the maximum possible DGD is known, it is useful to limit the length n of the initial impulse response matrix \mathbf{h}_n appropriately before inverse scattering is started.

The algorithm was tested experimentally. A tunable laser module (191.7...196.1 THz in 50 GHz steps) was connected to a polarization scrambler/transformer (Novoptel EPS1000), the device under test (DUT) and a 100 MS/s polarimeter (Novoptel PM1000), see Fig. 2. By applying several pre-determined polarization states, the Mueller matrix of the DUT was determined as a function of optical frequency.

To avoid detrimental artifacts of chromatic dispersion, the sign of a Jones matrix \mathbf{J} was inverted if the dominating

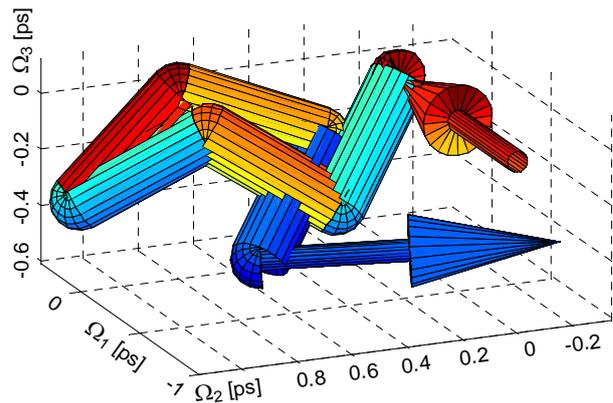


Fig. 5. DGD profile of 7.6 dB polarizer with 750 fs of DGD

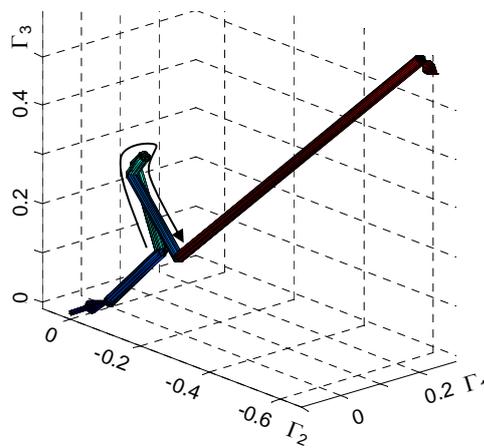


Fig. 6. PDL profile, containing extinction-based PDL vectors, of 7.6 dB polarizer with 750 fs of DGD. Not needed PDL sections, which fold up and thereby compensate each other pairwise, are marked by a curved arrow.

elements differed in phase by more than $\pi/2$ with respect to those of the neighboring Jones matrix. For the same reason, we set $\arg(\det(\mathbf{J}))$ to be the same for all \mathbf{J} . The corresponding Jones matrix spectrum was windowed, Fourier-backtransformed and 2-fold undersampled into the sequence of \mathbf{h}_i . Each DGD section was $\tau = 455$ fs long. This value equals the inverse of the total frequency scanning range, times 2 because of undersampling. The available number of sections is determined by the number of tuning steps, divided by 2 because of undersampling. The total DGD range (20 ps) equals the inverse of the tuning steps. After inverse scattering, input-referred DGD (Ω) and PDL (Γ) profiles of the DUT were plotted in the space of normalized Stokes vectors. These profiles contain, chained one after the other, the individual DGD or PDL vectors whose sum equal the total 1st-order PMD or approximately the total PDL, respectively. For accuracy reasons the extinction-based PDL vectors of Section II. are used. The angle between adjacent DGD or PDL sections represents the normalized Stokes vector rotation in the retarder separating the adjacent sections.

In the following Figs., note that adjacent PMD or PDL vectors with equal lengths but opposite orientations cancel each other. The initial impulse response matrix can be

truncated in length if it contains only negligibly small elements at its edges and the resulting structural profiles are too messy. The input arrow denotes horizontal input polarization, and the output arrow denotes that input polarization which would be needed, when PDL was neglected, to excite a principal state-of-polarization of the n th PDL section or of a fictitious additional $(n + 1)$ th DGD section.

Two polarization-maintaining fiber (PMF) pieces with 4 ps and 6.6 ps of DGD were concatenated under random orientation. The resulting DGD profile (with PMD vector components $\Omega_{1\dots3}$ in ps) is shown in Fig. 3. Apparent are the straight DGD profile portions correspondig to these two PMF pieces. Only about 8 (3.64 ps) + 14 (6.37 ps) = 22 DGD sections are needed for these. However, more DGD sections are delivered by the inverse scattering algorithm. Those which are not needed fold up fairly accurately (see curved arrows) so that they compensate each other pairwise or make up for the 4 ps rather than 3.64 ps DGD of one of the PMF pieces. The distance between the head of the input arrow and the tail of the output arrow is the total DGD. A reference measurement would show head of input and tail of output arrow glued together, and in between a double strand of DGD sections which cancel each other pairwise (see Fig. 13 of [9], without PDL considered).

Fig. 4 displays the associated PDL profile (with extinction-based PDL vector components $\Gamma_{1\dots3}$). PDL was set to zero for impulse response matrices with determinant magnitudes $< 2.5 \cdot 10^{-5}$. Ideally, one would expect a neutral picture with nothing to be seen except input and output arrows. This is not the case. However, most local PDL vector excursions, probably caused by measurement errors, are compensated by adjacent opposite ones, similar to the DGD sections near the output arrow of Fig. 3. Total (approximated) PDL vector length between the head of input arrow and the tail of output arrow is close to 0 as expected, $|\Gamma| = \gamma \approx 0.01$. Input/output arrow directions are identical for DGD and PDL profiles.

As another DUT, a piece of special PMF was chosen. When bent, it exhibited a PDL of 7.6 dB or $\gamma \approx 0.87$. DGD was 750 fs. To avoid profile overflow, we truncated the matrix impulse response to $n = 6$ before inverse scattering. Fig. 5 shows the DGD profile. The 2nd to 5th section mostly compensate each other. The total DGD between head of input and tail of output arrow was ≈ 750 fs as expected.

Fig. 6 is the corresponding PDL profile. Here, too, the 2nd to 5th section mainly compensate each other. The 1st and 6th section have the same direction, thereby making up for the full device PDL. The total PDL measured vector length is ≈ 0.87 as expected.

Closer inspection shows that the overall DGD vector in Fig. 5 and the overall PDL vector in Fig. 6 are not far from being parallel (because PDL and DGD have equal eigenmodes along the DUT) with opposite signs.

Note that the PDL of the special PMF depended somewhat on frequency. This could only be correctly represented by our

structural model if the PDL change and a similar total DGD change were brought about by the frequency-dependent phase delays in the DGD sections. That was not the case here. Rather, frequency-dependent mode stripping determined PDL.

IV. PDL MEASURED BY SQRT(3) SCRAMBLING METHOD

We want to measure the PDL of a device under test (DUT). It has the Mueller matrix \mathbf{P} with the elements P_{jk} ($j, k \in \{0,1,2,3\}$). Its output intensity is given by the 0th line of the Mueller matrix equation, $S_{0,o} = (P_{00} + \mathbf{p}^T \mathbf{S}_i) S_{0,i}$. Here, $S_{0\dots3}$ are Stokes parameters, $\mathbf{S} = [S_1 \ S_2 \ S_3]^T / S_0$ is a normalized Stokes vector, the additional indexes mean i for input, o for output, P_{00} is the mean power transmission and the linear, not normalized PDL vector [2] is $\mathbf{p} = [P_{01} \ P_{02} \ P_{03}]^T$. Here, we need to use a linear PDL vector because to get the variance of a quantity (= intensity within the limits $T_{\text{ga}} e^{\pm\gamma}$), the quantity itself and not something like its logarithm (γ) must be measured. Maximum and minimum output intensities are obtained in the cases $\mathbf{p} = \pm a \mathbf{S}_i$ where $a > 0$ is a proportionality constant. For fully polarized input light they assume the values $S_{0,o} = (P_{00} \pm |\mathbf{p}|) S_{0,i}$ with $|\mathbf{p}| = \sqrt{\mathbf{p}^T \mathbf{p}}$. The PDL in dB is

$$\text{given through PDL [dB]} = 10 \log_{10} \frac{P_{00} + |\mathbf{p}|}{P_{00} - |\mathbf{p}|} = 10 \log_{10} \frac{T_{\text{max}}}{T_{\text{min}}}$$

It is relatively easy to generate random polarization states of the input signal with a polarization scrambler. These states are needed for our PDL measurement method. The three elements of the normalized Stokes vector \mathbf{S}_i of the random input polarizations are equidistributed in the interval $[-1, 1]$. They have zero mean and are uncorrelated. This means \mathbf{S}_i is equidistributed on the Poincaré sphere; its covariance matrix is discussed below. For simplicity, we assume that $S_{0,i}$ and the elements of \mathbf{S}_i are uncorrelated. As a consequence, the output intensity $S_{0,o}$ features the expectation value $\langle S_{0,o} \rangle = P_{00} \cdot \langle S_{0,i} \rangle$ and the variance

$$\sigma_{S_{0,o}}^2 \approx P_{00}^2 \sigma_{S_{0,i}}^2 + \mathbf{p}^T \langle \mathbf{S}_i \mathbf{S}_i^T \rangle \mathbf{p} \langle S_{0,i} \rangle^2 \quad (35)$$

For constant $S_{0,i}$, i.e. $\sigma_{S_{0,i}} = 0$, the standard deviation is obtained as $\sigma_{S_{0,o}} = \sqrt{\mathbf{p}^T \langle \mathbf{S}_i \mathbf{S}_i^T \rangle \mathbf{p}} \cdot \langle S_{0,i} \rangle$. It is bounded by $|\mathbf{p}| \sqrt{s_{\text{min}}} \cdot \langle S_{0,i} \rangle \leq \sigma_{S_{0,o}} \leq |\mathbf{p}| \sqrt{s_{\text{max}}} \cdot \langle S_{0,i} \rangle$. Here, s_{min} , s_{max} are minimum and maximum eigenvalues of the covariance matrix $\langle \mathbf{S}_i \mathbf{S}_i^T \rangle$. Ideally, $\langle \mathbf{S}_i \mathbf{S}_i^T \rangle$ equals 1/3 times the identity matrix [29], its eigenvalues coincide as

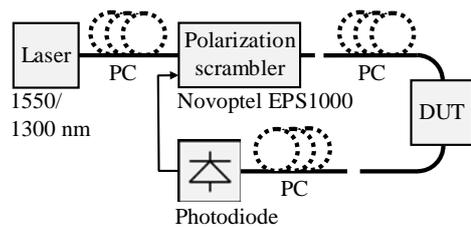


Fig. 7. Setup for PDL measurement with sqrt(3) scrambling and extinction methods. The dashed manual polarization controllers (PC) are not needed in practice; they just serve to assess measurement accuracy.

Waveplate	Rotations	Start at
QWP0	2^{10}	1/48
QWP1	2^6	3/48
QWP2	2^2	5/48
HWP	2^{12}	0
QWP3	2^8	7/48
QWP4	2^4	9/48
QWP5	2^0	11/48

Table 1. Waveplate rotations (in each 10.7 s long measurement) and starting positions (referred to 1 electrooptic revolution)

$s_{\min} = s_{\max} = 1/3$, and it holds $\sigma_{S_{0,o}} = |\mathbf{m}|/\sqrt{3} \cdot \langle S_{0,i} \rangle$. In the sqrt(3) scrambling method we calculate PDL in dB by
$$\text{PDL [dB]} = 10 \log_{10} \frac{\langle S_{0,o} \rangle + \sqrt{3} \sigma_{S_{0,o}}}{\langle S_{0,o} \rangle - \sqrt{3} \sigma_{S_{0,o}}}$$
. PDL values that

measured small in dB can fall in the range $\sqrt{3s_{\min}} \dots \sqrt{3s_{\max}}$ times the true PDL in dB. Minimum and maximum loss in dB are $10 \log_{10} \left(\frac{\langle S_{0,o} \rangle \pm \sqrt{3} \sigma_{S_{0,o}}}{\langle S_{0,i} \rangle} \right)$. Note that the $\sqrt{3}$ shows up already in eqn. (5) of [1], which points to the principle.

The sqrt(3) scrambling method has been implemented. The PDL of the polarization scrambler should be mathematically eliminated by a reference measurement. Simultaneous intensity measurement at the output and at the input of the DUT requires a splitter with extremely low PDL. We have avoided this need by sequentially measuring intensity with DUT in place, and without DUT for reference. An optomechanical fiber switch, having an extremely low PDL, can facilitate this but we simply use connectors to insert the DUT. The sequential measurement of DUT and reference

requires the PDL of the scrambler to be reproducible. In practice, the same polarizations must be used for DUT and reference measurements.

Fig. 7 shows the setup. A 1550 nm and a 1300 nm laser are available. The laser signal is passed through the polarization scrambler. This test signal is connected to a photodiode, either directly or through the DUT. The photodiode is built into, and connected to the scrambler, which in turn is controlled by a host computer. The photocurrent is sampled and converted to digital every 10 ns. For increased accuracy, conversion results are averaged over 164 μs (16384 samples). In order to assess measurement accuracy, manual polarization transformers are inserted before and behind the polarization scrambler and before the photodetector. These are not needed in practice.

For each measurement, the same sequence of 32768 scrambler settings is applied, one every 328 μs , and the resulting averaged intensity is measured in the second half of this interval. Measurement (10.7 s) and data transfer to the host need together 30 s. The 32768 test polarizations are generated by equispaced time-discrete rotation steps of 3 quarterwave plates (QWP), 1 halfwave plate (HWP) and 3 more QWPs in the electrooptic polarization scrambler (Table 1). The HWP rotates fastest and completes 4096 electrooptic rotation periods in 8 steps each. QWP5 rotates slowest and completes 1 period in 32768 steps. Rotation speeds differ by factors equal to powers of 4. All waveplates start from different orientations.

To start with, the 32768 test polarizations were recorded with our polarimeter. The degree-of-polarization of their mean was 0.015. The eigenvalues of the covariance matrix were computed without prior subtraction of the mean in order to obtain an upper limit for the fundamental error. The eigenvalues, 0.3294, 0.3313 and 0.3393, resulted in $\sqrt{s_{\max}/s_{\min}} = 1.0148$ and a fundamental $\pm 0.74\%$ relative error of small PDL values in dB. The polarimeter was used only to investigate the achievable accuracy, not for PDL measurement.

In total we have characterized 6 different PDL elements: a patchcord, a LiNbO₃ component, a short piece of special PMF (3M), which becomes polarizing when bent in 3 different bending states, and a polarizer (Table 2). Each measurement was repeated at least 10 times with different settings of the

PDL element	Sqrt(3) scrambling method		Extinction method	
	Minimum loss	PDL	Minimum loss	PDL
5m patchcord	0.0325 \pm 0.024 dB	0.0186 \pm 0.0091 dB	0.032 \pm 0.02 dB	<i>0.0104 \pm 0.021 dB</i>
LiNbO ₃ component (EOSPACE)	1.67 \pm 0.05 dB	0.0799 \pm 0.018 dB	1.69 \pm 0.055 dB	<i>0.0602 \pm 0.046 dB</i>
Weakly bent PMF (3M)	0.39 \pm 0.025 dB	0.433 \pm 0.041 dB	0.37 \pm 0.043 dB	<i>0.428 \pm 0.067 dB</i>
Moderately bent PMF (3M)	0.418 \pm 0.02 dB	1.59 \pm 0.087 dB	0.41 \pm 0.017 dB	1.6 \pm 0.091 dB
Strongly bent PMF (3M)	0.414 \pm 0.036 dB	7.3 \pm 0.074 dB	0.415 \pm 0.023 dB	7.33 \pm 0.051 dB
Polarizer	0.465 \pm 0.21 dB	29.85 ... ∞ dB	0.4 \pm 0.32 dB	50.4 \pm 1.5 dB

Table 2. Mean loss and PDL measurement results for different PDL elements. PDL results in italics are considered to be less accurate.

Signal wavelength	1550 nm		1300 nm	
Scaling parameter wavelength	1550 nm	1300 nm	1550 nm	1300 nm
PDL with sqrt(3) scrambling method	0.417 \pm 0.018 dB	0.416 \pm 0.019 dB	0.144 \pm 0.037 dB	0.15 \pm 0.037 dB

Table 3. Measured PDL of weakly bent special PMF (3M) at two wavelengths and with two voltage scaling parameters.

three manual polarization controllers (PC in Fig. 7), in order to show that results almost do not depend on scrambler input polarization and the particular polarization at and behind the DUT. The indicated error windows contain all these results while the associated standard deviations are smaller.

With the $\sqrt{3}$ scrambling method, residual PDL of the 5m patchcord, connector, manual polarization controllers and photodetector is < 0.02 dB. Even the small PDL of the LiNbO_3 component is measured very accurately with errors < 0.02 dB. The PDL error for the strongly bent PMF is 0.074 dB. In relation to the mean PDL of 7.3 dB this is 1%. Not surprisingly, this figure surpasses the fundamental relative PDL error of 0.074% derived above. The polarizer extinction is determined inaccurately because for the close-to-ideal polarizer it holds $\langle S_{0,o} \rangle \approx \sqrt{3}\sigma_{S_{0,o}}$, which results in strong error propagation.

The extinction method has likewise been implemented: The host computer dithers scrambler voltage settings around the operation point. Then the operation point is moved in or against the direction of the measured gradient. After many executed gradient steps either maximum or minimum transmission is found. From these, PDL is determined.

Also for the extinction method, a PDL reference is measured, using the scrambler voltage settings found for maximum and minimum transmission. This can reduce the influence of scrambler PDL. Furthermore, all measurements are repeated at least 10 times. At small PDL values, the extinction method cannot be considered as reliable because maximum and minimum transmissions are falsified by scrambler PDL. In Table 2, the results assumed to be less accurate (= having larger deviations from the mean) are printed in italics. The strength of the extinction method is the accurate measurement of large extinctions, even surpassing 50 dB for the investigated polarizer. Measured extinction was double-checked with a hand-held power meter.

Next, wavelength independence of the $\sqrt{3}$ scrambling method was assessed. Waveplate retardations scale with optical frequency. For 1300 nm rather than 1550 nm operation, LiNbO_3 electrode voltages are therefore decreased by a scaling parameter. That is equivalent to creating a different scrambler. With scaling parameters set for 1300 nm, the fundamental relative PDL error at 1550 nm (computed from s_{\max} , s_{\min}) was $\pm 0.7\%$, no worse than with 1550 nm scaling parameters. This is understood from the large number of waveplates (6 QWPs, 1 HWP) in the scrambler.

We have measured the PDL of the weakly bent special PMF at both wavelengths with both electrode voltage scaling parameters (Table 3). While PDL, as could be expected, differs considerably between 1300 nm and 1500 nm, the measurements at a given wavelength depend only slightly on the scaling parameter, thereby showing wavelength independence of the $\sqrt{3}$ scrambling method.

One could argue that this is no strict proof for the wavelength-independence of the PDL measurement. But is PDL measurement with a polarimeter more truthful in this respect? No, because polarimeters are usually calibrated at

various wavelengths using scrambled input polarizations!

The same holds for the scrambler characterization which gave the covariance matrix eigenvalues and fundamental relative error: The characterization, based on a polarimeter, is probably not more correct than our or any other scrambler. However, the different polarization settings undertaken to obtain the error intervals in Table 2 are equivalent to creating many different scramblers. The fact that the error intervals produced by these are small proves that the principle is sound and the measurements are accurate. We consider this proof at least as credible as the prediction of low fundamental relative error ($\pm 0.74\%$) by the polarimetric characterization.

V. DISCUSSION AND CONCLUSION

To our knowledge, the proposed usage of input- and output-referred PDL vectors based on extinction units is new, and so is the explicit specification of the rotation matrix $\tilde{\mathbf{W}}$ and the decomposition of a general constant optical element into a retarder, a PDL element and a geometric average power transmission. Our redefined, extinction-based PDL vectors can be added, in particular if they are not small, with much better accuracy than the traditional ones based on linear units.

This allows the plotting of PDL and DGD profiles as intuitive graphical displays of the distributed optical device structure. For this purpose, the device is modeled by alternating DGD and PDL sections and retarders. A Mueller matrix measurement in the frequency domain successively yields a Jones matrix spectrum, a Jones matrix impulse response and, by means of inverse scattering, the device structure.

The $\sqrt{3}$ scrambling method has been introduced for simple, accurate, low-cost PDL measurement. Even though no polarimeter is needed, this technique is independent of scrambler input polarization and wavelength. The advantage over the all-states method is that essentially all, not just the extreme intensity values enter into the calculation and make the result accurate. At low PDL values the $\sqrt{3}$ scrambling method is found to be more accurate than the gradient search based extinction method, which is known from polarization tracking experiments and has also been implemented.

REFERENCES

- [1] N. Gisin, "Statistics of polarization dependent losses", *Opt. Commun.*, 114(1995)5-6, pp. 399-405
- [2] C. Vinegoni, M. Karlsson, M. Petersson, H. Sunnerud, "The statistics of polarization-dependent loss in a recirculating loop", *IEEE J. Lightwave Technology*, 22(2004)4, pp. 968 – 976
- [3] A. Mecozzi, M. Shtaf, "The Statistics of Polarization-Dependent loss in Optical Communication Systems", *IEEE PTL*, 14(2002)3, pp. 313 – 315
- [4] M. Karlsson, M Petersson, "Quaternion Approach to PMD and PDL Phenomena in Optical Fiber Systems", *IEEE JLT*, 22(2004)4, pp. 1137 – 1146
- [5] N. Gisin, B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers", *Opt. Commun.*, 142(1997), pp. 119-125
- [6] B. Huttner, C. Geiser, N. Gisin, "Polarization-Induced Distortions in Optical Fiber Networks with Polarization-Mode Dispersion and Polarization-Dependent Losses", *IEEE JSTQE*, 6(2000)2, pp. 317 – 329

- [7] Y. Li, A. Yariv, "Solutions to the dynamical equation of polarization-mode dispersion and polarization-dependent losses", *JOSA B*, 17(2000)11, pp. 1821-1827
- [8] L. E. Nelson et al. "Statistics of polarization dependent loss in an installed long-haul WDM system," *Optics Express*, Vol. 19, No. 7, pp. 6790-6796, March 2011
- [9] D. Sandel et al., "Some enabling techniques for polarization mode dispersion compensation," *J. Lightwave Technol.*, Vol. 21, no. 5, pp. 1198-1210 (2003)
- [10] R. Noé et al., "PMD in High-Bitrate Transmission and Means for its Mitigation." *J. Selected Topics in Quantum Electronics*, Vol. 10, no. 2, pp. 341-355 (2004)
- [11] S.E. Harris, E.O. Ammann, I.C. Chang, *Optical Network Synthesis Using Birefringent Crystals. I. Synthesis of Lossless Networks of Equal-Length Crystals.* *J. Opt. Soc. Am.* 54 (1964), pp. 1267-1279
- [12] L. Möller, *Filter Synthesis for Broad-Band PMD Compensation in WDM Systems.* *IEEE Photonics Technol. Lett.* 12(2000), pp. 1258-1260
- [13] IEC61300-3-12
- [14] System and Method for measuring Polarization Dependent Loss, D.L. Favin, B.M. Nyman, G.M. Wolter, US5371597, 1994
- [15] B. Heffner, Deterministic, analytically complete measurement of polarization-dependent transmission through optical devices, *IEEE Photonics Technology Letters*, 4(1992)5, pp. 451-454
- [16] B. Nymann, D. Favin, G. Wolter, Automated system for measuring polarization dependent loss, in *Proc. OFC 1994*, 20 Feb. 1994, San Jose, CA, ThK6, pp. 230-231
- [17] R. Craig, S. Gilbert, P. Hale, High-resolution, nonmechanical approach to polarization dependent transmission measurements, *IEEE J. Lightwave Technol.* 16(1998)7, pp. 1285-1294
- [18] R. Craig, Accurate spectral characterization of polarization-dependent loss, *IEEE J. Lightwave Technol.* 21(2003)2, pp. 432-437
- [19] <http://cp.literature.agilent.com/litweb/pdf/5990-9973EN.pdf>
- [20] C. Hentschel, D. Derickson, "Insertion loss measurements," in *Fiber Optic Test and Measurement*, D. Derickson, Ed., Upper Saddle River, NJ: Prentice-Hall, 1998, ch. 9
- [21] Y. Shi, L. Yan, X.S. Yao, Automatic Maximum-Minimum Search Method for Accurate PDL and DOP Characterization, *IEEE JLT* 24(2006)11, pp. 4006-4012
- [22] US020060115199A1
- [23] Noé, R., Heidrich, H., Hoffmann, D., Endless polarization control systems for coherent optics, *IEEE J. Lightwave Technol.* 6(1988)7, pp. 1199-1207
- [24] R. Noé, *Essentials of Modern Optical Fiber Communication*, Springer Berlin Heidelberg, March 2010, DOI 10.1007/978-3-642-04872-2
- [25] C. Antonelli et al. "Autocorrelation of the polarization-dependent loss in fiber routes," *Optics Letters*, Vol. 36, No. 20, pp. 4005-4007, October 2011
- [26] G.J. Foschini, C.D. Poole, Statistical theory of polarization dispersion in single mode fibers, *IEEE J. Lightwave Technol.* 9(1991)11, pp. 1439-1456
- [27] R. Noé et al., Polarization mode dispersion compensation at 10, 20 and 40 Gb/s with various optical equalizers. *IEEE J. Lightwave Technology*, 17(1999)9, pp. 1602-1616
- [28] H. Kogelnik, P.J. Winzer, L.E. Nelson, R.M. Jopson, M. Boroditsky, M. Brodsky, First-Order PMD Outage for the Hinge Model, *IEEE PTL* 17(2005)6, pp. 1208-1210
- [29] R. Noé et al., Polarization mode dispersion detected by arrival time measurement of polarization-scrambled light. *IEEE J. Lightwave Technol.* 20(2002)2, p. 229

Reinhold Noé (M'93) was born in Darmstadt, Germany, in 1960. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from Technische Universität München (Munich), Germany, in 1984 and 1987, respectively. At that time, he realized the first endless polarization control systems. Then he spent a postdoctoral year with Bellcore, Red Bank, NJ, where he continued to work on coherent optical systems. In 1988 he joined Siemens Research Laboratories, Munich. Since 1992, he is Chair of Optical Communication and High-Frequency Engineering, University of Paderborn, Germany. Together with Benjamin Koch he founded in 2010 Novoptel GmbH, Paderborn, Germany. Most of his recent experiments deal with high-speed endless polarization control and realtime synchronous QPSK transmission. He has (co)authored about 300 journal and conference publications.

Benjamin Koch was born in Gutweiler, Germany in 1977. He received his Dipl.-Ing. degree in electrical engineering from the University of Paderborn,

Germany, in 2007. Since then he is with the research group of Prof. Noé. He obtained his doctorate (Dr.-Ing.) in 2012. In 2010 he founded, together with Reinhold Noé, Novoptel GmbH, Paderborn, Germany, where he serves as Managing Director. His research concentrates on optical polarization control systems. He has (co-)authored over 30 peer-reviewed journal and conference papers.

David Sandel was born in Wangen/Allgäu, Germany, on December 26, 1966. He received the Dipl.-Ing degree from the Technische Hochschule Karlsruhe, Germany, in 1992 and the Dr.-Ing. degree from the University of Paderborn, Paderborn, Germany, in 1997. In the same year, he joined the Department of Optical Communication and High-Frequency Engineering, University of Paderborn, Germany. His doctoral work was on fiber Bragg gratings. His research interest now concentrates on polarization-mode dispersion compensation and polarization multiplex data transmission.

Vitali Mirvoda was born in Taganrog, Russia, in 1974. He received the M.S. degree in electrical engineering from Taganrog State University of Radio Engineering, Taganrog, Russia, in 1996 and the Dr.-Ing. degree from the University of Paderborn, Paderborn, Germany, in 2008. Since 1998 he has been engaged in research and engineering of optical communication systems at the University of Paderborn. He is currently a postdoctoral researcher with the Department of Optical Communication and High-Frequency Engineering. His research areas include polarization-mode dispersion compensation and chromatic dispersion.