Available online at www.sciencedirect.com
ScienceDirect
Optics Communications 274 (2007) 116–123
OPTICS COMMUNICATIONS
www.elsevier.com/locate/optcom

Measurement of Mueller matrix for an optical fiber system with birefringence and polarization-dependent loss or gain

H. Dong [a,*], P. Shum [a], M. Yan [a], J.Q. Zhou [a], G.X. Ning [a], Y.D. Gong [b], C.Q. Wu [c]

[a] Network Technology Research Centre, Nanyang Technological University, Singapore 637553, Singapore
[b] Institute for Inforcomm Research, Singapore 119613, Singapore
[c] School of Science, Beijing Jiaotong University, Beijing 100044, PR China

Received 8 February 2006; received in revised form 5 November 2006; accepted 8 February 2007

Abstract

In an optical fiber system with both birefringence and polarization-dependent loss or gain (PDL/G), a set of input polarization states and their corresponding outputs are deduced to fulfill some general relationships in Stokes space, by considering the fact that the Mueller matrix of such an optical fiber system meets Lorentz transformation. Then, it can be proven that the minimum number of input polarization states is three for an explicit determination of such a Mueller matrix and no independent input parameter is redundant. Based on the theoretical findings, a general and simple approach is proposed to measure the Mueller matrix. The requirements regarding the choices of three inputs are presented for achieving optimum measurement stability and accuracy. Experimental results on an optical fiber system, formed by two 10-km long single-mode fibers with an optical isolator in between, show excellent agreement with the theory.

Keywords: Optical fiber polarization; Mueller matrix; Birefringence; Polarization dependent loss or gain

1. Introduction

Polarization effects, such as polarization-dependent loss or gain (PDL/G) and polarization mode dispersion (PMD) are demonstrated to have significant impact on an optical fiber system [1,2]. As we know, polarization state of a light is completely described by 4 Stokes parameters, and polarization properties of a non-image-forming optical system can be completely represented by a 4×4 Mueller matrix which relates the input and output Stokes parameters of light [3]. Measurement of Mueller matrix for an optical fiber system is therefore of high importance for the investigation of PMD and PDL/G in optical fibers. In an optical system with pure birefringence, Mueller matrix only has 3 degrees of freedom (DOFs), which represent three rotations in Stokes space. Optical power and degree of polarization (DOP) are not changed for a light passing through such a system. Further, it has been noticed that the dot product of two 3-dimensional (3D) Stokes vectors and the scalar triple product (representing 3D volume) of three 3D Stokes vectors are conservative in a pure birefringent optical system. The conservation of the dot product as well as 3D volume has been used to measure Mueller matrices of single-mode optical fibers employing two input polarization states, and in turn to determine the first- and second-order PMD vectors [4]. In the most general case, 16 elements in the 4×4 Mueller matrix are independent. Such a Mueller matrix includes three basic polarization effects: birefringence (retardance), PDL/G (diattenuation) and depolarization [5]. It is very difficult, if not impossible, to find any relationships between input and output Stokes parameters for a system with 16 DOFs. So four input polarization states (corresponding to 16 independent parameters) have to be used to measure such a Mueller matrix. For an optical fiber system with both birefringence and PDL/G, it has been pointed out that there are 7 DOFs in such a system, including three rotations (corresponding

* Corresponding author.
E-mail address: hdong@ntu.edu.sg (H. Dong).

0030-4018/$ - see front matter
doi:10.1016/j.optcom.2007.02.015

to birefringence), three boost actions (corresponding to PDL/G) and one length compression/expansion (corresponding to attenuation/gain for unpolarized light) [5]. It can be verified that in such a system the optical power and DOP (unless DOP = 1) are no longer conservative [6]. For the determination of transmission matrix in such an optical system, R.C. Jones gave a method for measurement of the Jones matrix in 1940s using three well-defined input polarization states $(1\ 0\ 0)^{\mathrm{T}}$, $(-1\ 0\ 0)^{\mathrm{T}}$ and $(0\ 1\ 0)^{\mathrm{T}}$ [7], and the Jones matrix can be then converted to Mueller matrix based on their relations [3]. Recently we proposed a method to measure Mueller matrix of such an optical fiber system directly in Stokes space also using three above-mentioned well-defined input polarization states [8]. However, in [8], the choices of three inputs were not optimized for achieving the optimum measurement accuracy as well as measurement stability against various environmental perturbations. To solve this problem, in this paper, we first derive some relationships between a set of input polarization states and their corresponding outputs for such a system based on the Lorentz Group property of Mueller matrix. Based on these relationships, it is proven that three input polarization states are sufficient but not redundant for the measurement of Mueller matrix. Then, we propose a general method employing three arbitrary input polarization states to measure the Mueller matrix. The conditions on the three inputs under which the method is stable and has the optimum accuracy are presented in the Appendix. Experimental results on an optical fiber system composed of single-mode fibers and optical isolator confirm the validity of our theoretical findings.

2. Relationships between input polarization states and their corresponding outputs

It has been pointed out theoretically and demonstrated experimentally that Mueller matrix of an optical system with both birefringence and PDL/G satisfies the Lorentz transformation, i.e., Lorentz Group O(3,1) [9,10]. It is therefore convenient to discuss the polarization issues in the 4-dimensional (4D) Minkowski space. Much like those in the Special Relativity [11], the 4 Stokes parameters are rewritten as a 4D complex vector $\vec{S} = (\mathrm{i}s_0\ s_1\ s_2\ s_3)^{\mathrm{T}}$, where $\mathrm{i} = \sqrt{-1}$ and superscript T denotes matrix transposition. Complex number is used as a complex matrix is more appropriate to describe a Lorentz transformation. Then the traditional 4×4 Mueller matrix $\mathbf{M}$ is rewritten as

$$\widetilde{\mathbf{M}} = \begin{pmatrix} m_{11} & \mathrm{i}m_{12} & \mathrm{i}m_{13} & \mathrm{i}m_{14} \\ -\mathrm{i}m_{21} & m_{22} & m_{23} & m_{24} \\ -\mathrm{i}m_{31} & m_{32} & m_{33} & m_{34} \\ -\mathrm{i}m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix}. \tag{1}$$

It can be noted that $|\widetilde{\mathbf{M}}| = |\mathbf{M}|$, where $|\cdot|$ denotes the determinant of a square matrix or the magnitude of a 3D vector. Based on this expression, new Mueller matrix $\widetilde{\mathbf{M}}$, as a direct property of Lorentz Group, should meet [10,12]

$$\widetilde{\mathbf{M}}^{\mathrm{T}}\widetilde{\mathbf{M}} = \sqrt{|\widetilde{\mathbf{M}}|}\mathbf{I}. \tag{2}$$

Here $\mathbf{I}$ is an 4×4 identity matrix.

From Eq. (2), two pairs of inputs and outputs $\vec{S}_{\mathrm{out}} = \widetilde{\mathbf{M}}\vec{S}_{\mathrm{in}}$ and $\vec{T}_{\mathrm{out}} = \widetilde{\mathbf{M}}\vec{T}_{\mathrm{in}}$ should fulfill

$$\vec{S}_{\mathrm{out}} \cdot \vec{T}_{\mathrm{out}} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{S}_{\mathrm{in}} \cdot \vec{T}_{\mathrm{in}}. \tag{3}$$

A simple case of Eq. (3) is [12]

$$\vec{S}_{\mathrm{out}} \cdot \vec{S}_{\mathrm{out}} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{S}_{\mathrm{in}} \cdot \vec{S}_{\mathrm{in}}. \tag{4}$$

In an optical system where only birefringence exists, the scalar triple product of three 3D vectors (3D volume) is conservative, which describes the handedness (chirality) of the system. In analogy, a factor that denotes the handedness of an optical system with both birefringence and PDL/G should be found out. Similar to the definition of scalar triple product, the scalar quadruple product (4D volume) of four 4D Stokes vectors $\vec{S}, \vec{T}, \vec{U}, \vec{V}$ is defined as

$$\mathrm{Vol}_{4\mathrm{D}} = \begin{vmatrix} \mathrm{i}s_0 & s_1 & s_2 & s_3 \\ \mathrm{i}t_0 & t_1 & t_2 & t_3 \\ \mathrm{i}u_0 & u_1 & u_2 & u_3 \\ \mathrm{i}v_0 & v_1 & v_2 & v_3 \end{vmatrix}. \tag{5}$$

Based on Eq. (2), it can be demonstrated that

$$\mathrm{Vol}_{4\mathrm{D}}^{\mathrm{out}} = |\widetilde{\mathbf{M}}|\mathrm{Vol}_{4\mathrm{D}}^{\mathrm{in}}. \tag{6}$$

3. The minimum number of input polarization states for determination of a Mueller matrix

In [8], it has been known that three well-defined input polarization states are enough to measure Mueller matrix. But it is still questionable if three arbitrary input polarization states are sufficient. And if the answer is yes, is any of them redundant? The answer depends on the relationships which have been derived in Section 2. In this section, we strictly prove that three arbitrary inputs are sufficient to measure Mueller matrix and no input parameters are redundant.

A 4D Stokes vector should have 4 independent parameters, thus 3 inputs generate 12 DOFs in general. But some DOFs will be cancelled out due to symmetric property of the optical system. Assuming there are three inputs $\vec{S}_{\mathrm{in}}$, $\vec{T}_{\mathrm{in}}$, $\vec{U}_{\mathrm{in}}$ and their corresponding outputs $\vec{S}_{\mathrm{out}}$, $\vec{T}_{\mathrm{out}}$, $\vec{U}_{\mathrm{out}}$, based on Eq. (3), three equations are obtained as

$$\begin{cases} \vec{S}_{\mathrm{out}} \cdot \vec{T}_{\mathrm{out}} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{S}_{\mathrm{in}} \cdot \vec{T}_{\mathrm{in}}, \\ \vec{S}_{\mathrm{out}} \cdot \vec{U}_{\mathrm{out}} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{S}_{\mathrm{in}} \cdot \vec{U}_{\mathrm{in}}, \\ \vec{U}_{\mathrm{out}} \cdot \vec{T}_{\mathrm{out}} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{U}_{\mathrm{in}} \cdot \vec{T}_{\mathrm{in}}. \end{cases} \tag{7}$$

Then three Mueller-matrix-independent equations in turn are deduced from Eq. (7)

$$\begin{cases} (\vec{S}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{U}_{\rm out}),\\ (\vec{U}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{U}_{\rm out}\cdot\vec{T}_{\rm out}),\\ (\vec{U}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{U}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{U}_{\rm out}\cdot\vec{T}_{\rm out}). \end{cases} \tag{8}$$

In Eq. (8), the third equation can be derived by using others, hence only two equations are independent. This implies that there are only 10 DOFs left now. In the remaining part of this section, we will consider four combinations of three input polarization states.

(1) *Three completely polarized inputs*
It has been known that the output is also completely polarized for an optical system with both birefringence and PDL/G subjecting to a completely polarized input [6]. This means that 3 DOFs will not take effect among three inputs; therefore only 7 DOFs exist for three completely polarized inputs. As has been mentioned in Section 1, the Mueller matrix has also 7 DOFs. Hence the Mueller matrix can be determined explicitly with three completely polarized inputs and there are no any redundant independent input parameters.

(2) *Two completely polarized and one partially polarized (including unpolarized) inputs*
For two completely polarized inputs, 2 DOFs are cancelled out and hence only 8 DOFs are left for this combination of inputs. Assuming $\vec{S}_{\rm in}$ is the partially polarized input, from Eqs. (4) and (7), two Mueller-matrix-independent equations are formed

$$\begin{cases} (\vec{S}_{\rm in}\cdot\vec{S}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{S}_{\rm out}),\\ (\vec{S}_{\rm in}\cdot\vec{S}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{U}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{S}_{\rm out}). \end{cases} \tag{9}$$

Only one equation is independent in Eq. (9), which may be confirmed by comparing Eq. (9) and the first equation of Eq. (8). Effectively this reduces the number of DOF to 7.

(3) *One completely polarized and two partially polarized (including unpolarized) inputs*
For one completely polarized input, 1 DOF is cancelled out and then 9 DOFs are left for three inputs. Assuming $\vec{S}_{\rm in}$ and $\vec{T}_{\rm in}$ are the partially polarized inputs, also from Eqs. (4) and (7), there are four Mueller-matrix-independent equations

$$\begin{cases} (\vec{S}_{\rm in}\cdot\vec{S}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{S}_{\rm out}),\\ (\vec{S}_{\rm in}\cdot\vec{S}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{U}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{S}_{\rm out}),\\ (\vec{T}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{T}_{\rm out}\cdot\vec{T}_{\rm out}),\\ (\vec{T}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{T}_{\rm out}\cdot\vec{U}_{\rm out}) = (\vec{T}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{T}_{\rm out}\cdot\vec{T}_{\rm out}). \end{cases} \tag{10}$$

Also by comparing with Eq. (8), only two equations in Eq. (10) are independent. That is, Eq. (10) cancels out two DOFs. Again, 7 DOFs are left for this combination of inputs.

(4) *Three partially polarized (including unpolarized) inputs*
Six Mueller-matrix-independent equations can be obtained from Eqs. (4) and (7) in this case as

$$\begin{cases} (\vec{S}_{\rm in}\cdot\vec{S}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{S}_{\rm out}),\\ (\vec{S}_{\rm in}\cdot\vec{S}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{U}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{S}_{\rm out}),\\ (\vec{T}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{S}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{S}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{T}_{\rm out}\cdot\vec{T}_{\rm out}),\\ (\vec{T}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{T}_{\rm out}\cdot\vec{U}_{\rm out}) = (\vec{T}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{T}_{\rm out}\cdot\vec{T}_{\rm out}),\\ (\vec{U}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{U}_{\rm out}\cdot\vec{S}_{\rm out}) = (\vec{U}_{\rm in}\cdot\vec{S}_{\rm in})(\vec{U}_{\rm out}\cdot\vec{U}_{\rm out}),\\ (\vec{U}_{\rm in}\cdot\vec{U}_{\rm in})(\vec{U}_{\rm out}\cdot\vec{T}_{\rm out}) = (\vec{U}_{\rm in}\cdot\vec{T}_{\rm in})(\vec{U}_{\rm out}\cdot\vec{U}_{\rm out}). \end{cases} \tag{11}$$

By comparing with Eq. (8), three equations in Eq. (11) are found to be independent. So Eq. (11) cancels out three DOFs, which also results in 7 DOFs.

From the above analysis, one conclusion can be drawn that three arbitrary input polarization states are sufficient but not redundant for the determination of Mueller matrix. In fact, it also can be demonstrated that two input polarization states only generate six independent input parameters in such an optical system by following a similar analysis. Hence we can not use only two input polarization states to determine the Mueller matrix in an optical fiber link with PDL/G.

4. Measurement of Mueller matrix

As demonstrated in Section 3, at least three inputs have to be used to measure Mueller matrix, and seven independent equations can be obtained from three input–output pairs. Apart from these, nine independent bilinear equations have been proposed to express relations of Mueller matrix elements based on Mueller matrix's property [12]. Then an equation group consisting of 16 independent equations can be solved to determine all elements of the Mueller matrix in principle. However, such method will lead to complicated calculations since some equations are nonlinear [12]. To better tackle this problem, three arbitrary inputs $\vec{S}_{\rm in}$, $\vec{T}_{\rm in}$, $\vec{U}_{\rm in}$ are launched and their corresponding outputs $\vec{S}_{\rm out}$, $\vec{T}_{\rm out}$, $\vec{U}_{\rm out}$ are measured in the first step. Then we can calculate (not measure) any other output $\vec{V}_{\rm out}$ corresponding to an arbitrarily assumed input $\vec{V}_{\rm in}$ according to the following steps:

(1) By using Eq. (3) and two measured input–output pairs, $\sqrt{|\widetilde{\mathbf{M}}|}$ is obtained;
(2) From Eqs. (3) and (6), four linear independent equations are obtained as

$$\begin{cases} \vec{S}_{\rm out}\cdot\vec{V}_{\rm out} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{S}_{\rm in}\cdot\vec{V}_{\rm in},\\ \vec{T}_{\rm out}\cdot\vec{V}_{\rm out} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{T}_{\rm in}\cdot\vec{V}_{\rm in},\\ \vec{U}_{\rm out}\cdot\vec{V}_{\rm out} = \sqrt{|\widetilde{\mathbf{M}}|}\vec{U}_{\rm in}\cdot\vec{V}_{\rm in},\\ {\rm Vol}^{\rm out}_{\rm 4D} = |\widetilde{\mathbf{M}}|{\rm Vol}^{\rm in}_{\rm 4D}. \end{cases} \tag{12}$$

By solving Eq. (12), $\vec{V}_{\rm out}$ can be calculated. The condition that Eq. (12) can be solvable is

$$\begin{vmatrix} \mathrm{i}s_{\mathrm{out0}} & s_{\mathrm{out1}} & s_{\mathrm{out2}} & s_{\mathrm{out3}} \\ \mathrm{i}t_{\mathrm{out0}} & t_{\mathrm{out1}} & t_{\mathrm{out2}} & t_{\mathrm{out3}} \\ \mathrm{i}u_{\mathrm{out0}} & u_{\mathrm{out1}} & u_{\mathrm{out2}} & u_{\mathrm{out3}} \\ A_{\mathrm{out0}} & A_{\mathrm{out1}} & A_{\mathrm{out2}} & A_{\mathrm{out3}} \end{vmatrix} = A_{\mathrm{out0}}^2 + A_{\mathrm{out1}}^2 + A_{\mathrm{out2}}^2 + A_{\mathrm{out3}}^2 \neq 0. \tag{13}$$

Here

$$A_{\mathrm{out0}} = -\begin{vmatrix} s_{\mathrm{out1}} & s_{\mathrm{out2}} & s_{\mathrm{out3}} \\ t_{\mathrm{out1}} & t_{\mathrm{out2}} & t_{\mathrm{out3}} \\ u_{\mathrm{out1}} & u_{\mathrm{out2}} & u_{\mathrm{out3}} \end{vmatrix}, \quad A_{\mathrm{out1}} = \begin{vmatrix} \mathrm{i}s_{\mathrm{out0}} & s_{\mathrm{out2}} & s_{\mathrm{out3}} \\ \mathrm{i}t_{\mathrm{out0}} & t_{\mathrm{out2}} & t_{\mathrm{out3}} \\ \mathrm{i}u_{\mathrm{out0}} & u_{\mathrm{out2}} & u_{\mathrm{out3}} \end{vmatrix}, \quad A_{\mathrm{out2}} = -\begin{vmatrix} \mathrm{i}s_{\mathrm{out0}} & s_{\mathrm{out1}} & s_{\mathrm{out3}} \\ \mathrm{i}t_{\mathrm{out0}} & t_{\mathrm{out1}} & t_{\mathrm{out3}} \\ \mathrm{i}u_{\mathrm{out0}} & u_{\mathrm{out1}} & u_{\mathrm{out3}} \end{vmatrix} \text{ and } A_{\mathrm{out3}} = \begin{vmatrix} \mathrm{i}s_{\mathrm{out0}} & s_{\mathrm{out1}} & s_{\mathrm{out2}} \\ \mathrm{i}t_{\mathrm{out0}} & t_{\mathrm{out1}} & t_{\mathrm{out2}} \\ \mathrm{i}u_{\mathrm{out0}} & u_{\mathrm{out1}} & u_{\mathrm{out2}} \end{vmatrix}.$$

This is the direct requirement regarding the output polarization states. As for the input polarization states, we can prove in Appendix that: (1) Eq. (12) can be solved when the three 3D input Stokes vectors are not linearly superposed in Stokes space if all inputs are completely polarized; (2) The optimum accuracy of Eq. (12) will be achieved when the three 3D input Stokes vectors are coplanar, the angles between each other are 120° and all inputs are completely polarized.

The above discussion means that if three input–output pairs are known by measurement, any other input–output pair can be calculated. Because all equations in Eq. (12) are *linear*, the calculation will be simple. Previously, people has used four *pre-determined* inputs to measure Mueller matrix, which are $I_0(1\ 1\ 0\ 0)^{\mathrm{T}}$, $I_0(1\ -1\ 0\ 0)^{\mathrm{T}}$, $I_0(1\ 0\ 1\ 0)^{\mathrm{T}}$ and $I_0(1\ 0\ 0\ 1)^{\mathrm{T}}$ (I_0 is the input optical power) [13]. In our method, firstly, three outputs $\vec{S}_{\mathrm{out}}$, $\vec{T}_{\mathrm{out}}$, $\vec{U}_{\mathrm{out}}$ corresponding to three inputs $\vec{S}_{\mathrm{in}}$, $\vec{T}_{\mathrm{in}}$, $\vec{U}_{\mathrm{in}}$ are measured. Then four outputs $I_1(1\,a_1\,a_2\,a_3)^{\mathrm{T}}$, $I_2(1\,b_1\,b_2\,b_3)^{\mathrm{T}}$, $I_3(1\,c_1\,c_2\,c_3)^{\mathrm{T}}$ and $I_4(1\,d_1\,d_2\,d_3)^{\mathrm{T}}$ corresponding to the above-mentioned four *pre-determined* inputs are calculated using Eq. (12). Finally the Mueller matrix under investigation can be achieved using the following formula

$$\mathbf{M} = \begin{pmatrix} \frac{1}{2}\left(\frac{I_1}{I_0}+\frac{I_2}{I_0}\right) & \frac{1}{2}\left(\frac{I_1}{I_0}-\frac{I_2}{I_0}\right) & \frac{I_3}{I_0}-\frac{1}{2}\left(\frac{I_1}{I_0}+\frac{I_2}{I_0}\right) & \frac{I_4}{I_0}-\frac{1}{2}\left(\frac{I_1}{I_0}+\frac{I_2}{I_0}\right) \\ \frac{1}{2}\left(\frac{a_1I_1}{I_0}+\frac{b_1I_2}{I_0}\right) & \frac{1}{2}\left(\frac{a_1I_1}{I_0}-\frac{b_1I_2}{I_0}\right) & \frac{c_1I_3}{I_0}-\frac{1}{2}\left(\frac{a_1I_1}{I_0}+\frac{b_1I_2}{I_0}\right) & \frac{d_1I_4}{I_0}-\frac{1}{2}\left(\frac{a_1I_1}{I_0}+\frac{b_1I_2}{I_0}\right) \\ \frac{1}{2}\left(\frac{a_2I_1}{I_0}+\frac{b_2I_2}{I_0}\right) & \frac{1}{2}\left(\frac{a_2I_1}{I_0}-\frac{b_2I_2}{I_0}\right) & \frac{c_2I_3}{I_0}-\frac{1}{2}\left(\frac{a_2I_1}{I_0}+\frac{b_2I_2}{I_0}\right) & \frac{d_2I_4}{I_0}-\frac{1}{2}\left(\frac{a_2I_1}{I_0}+\frac{b_2I_2}{I_0}\right) \\ \frac{1}{2}\left(\frac{a_3I_1}{I_0}+\frac{b_3I_2}{I_0}\right) & \frac{1}{2}\left(\frac{a_3I_1}{I_0}-\frac{b_3I_2}{I_0}\right) & \frac{c_3I_3}{I_0}-\frac{1}{2}\left(\frac{a_3I_1}{I_0}+\frac{b_3I_2}{I_0}\right) & \frac{d_3I_4}{I_0}-\frac{1}{2}\left(\frac{a_3I_1}{I_0}+\frac{b_3I_2}{I_0}\right) \end{pmatrix}. \tag{14}$$

This is a general approach, which can employ three arbitrary inputs in principle. But its accuracy is affected by various errors induced by equipments and fiber system under test. The optimum accuracy depends on the choice of three inputs in real measurements. The accuracy of Eq. (14) completely relies on the accuracy of four calculated outputs corresponding to four *pre-determined* inputs; and finally in turn totally depends on the choice of $\vec{S}_{\mathrm{in}}$, $\vec{T}_{\mathrm{in}}$, $\vec{U}_{\mathrm{in}}$ in Eq. (12). We have derived in the Appendix the conditions for solving Eq. (12) and also for achieving the optimum accuracy. It should be noticed that the conditions are also the requirements for the overall Mueller matrix method to be valid.

5. Experimental setup and results

To verify the proposed theory and measurement method, an experimental setup is established as shown in Fig. 1. The fiber system under test (FUT) is composed of two 10-km single-mode fibers (SMF), with an optical isolator in between. Such system has both birefringence (in SMF) and PDL (in optical isolator). The tunable laser source emits completely polarized light with stable polarization state as well as 700 kHz linewidth. The computer-controlled polarization controller (PC) as well as the in-line polarimeter are used to generate and monitor the appropriate input Stokes parameters. The polarimeter in the end is used to measure the output Stokes parameters. Two polarimeters we used are Thorlabs IPM5300 with polarization state accuracy of ±0.25° on Poincaré sphere and relative power measurement accuracy of ±0.01 dB. All components except FUT are controlled by a computer through GPIB port, USB port and I/O card. A measurement software is also developed based on this configuration.

To verify the above-mentioned theory, the experiments are implemented in two steps. Firstly four *pre-determined* inputs $I_0(1\,1\,0\,0)^{\mathrm{T}}$, $I_0(1\,-1\,0\,0)^{\mathrm{T}}$, $I_0(1\,0\,1\,0)^{\mathrm{T}}$, $I_0(1\,0\,0\,1)^{\mathrm{T}}$ ($I_0 = 2$ mW in our experiments) are generated one by one using the PC and in-line polarimeter and the corresponding

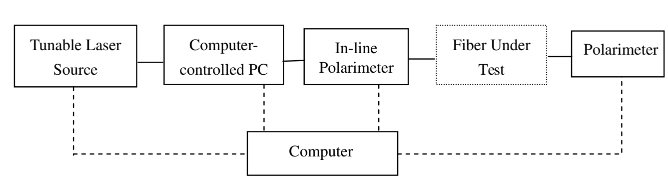

Fig. 1. Experimental configuration for Mueller matrix measurement in an optical fiber system.

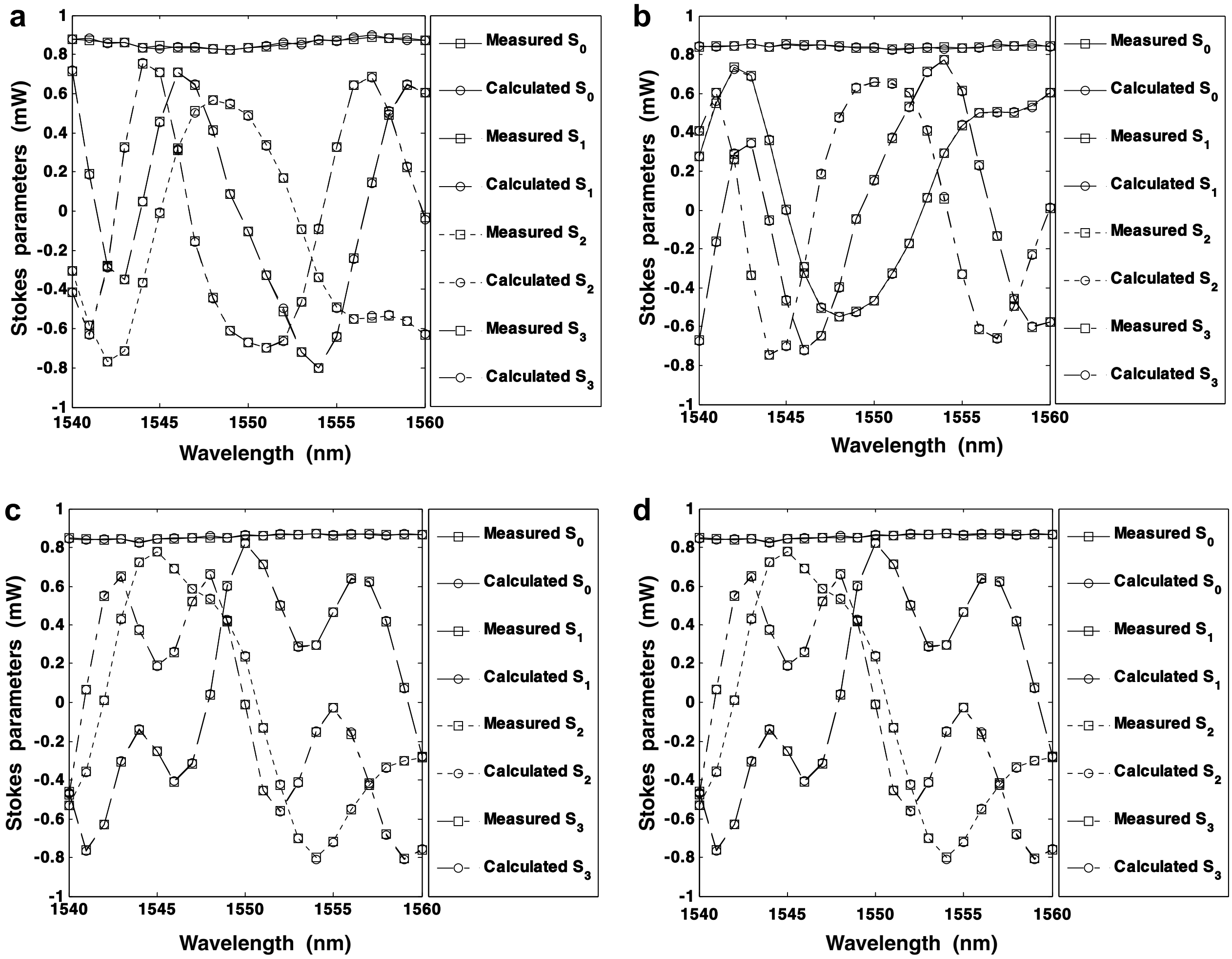

Fig. 2. Measured and calculated output Stokes parameters corresponding to four *pre-determined* inputs: (a) input polarization state (1 1 0 0); (b) input polarization state (1 −1 0 0); (c) input polarization state (1 0 1 0); (d) input polarization state (1 0 0 1).

outputs are measured as the benchmark. In the second step, according to the requirements for the optimum accuracy, three inputs $I_0(1\ 1\ 0\ 0)^{\mathrm{T}}$, $I_0(1\ \ -0.5\ 0.866\ 0)^{\mathrm{T}}$ and $I_0(1\ \ -0.5\ \ -0.866\ 0)^{\mathrm{T}}$ are generated one by one, also using the PC and in-line polarimeter, and the corresponding outputs are measured. Then the theoretical outputs corresponding to the above four *pre-determined* inputs are calculated from the measured data of three input–output pairs one by one using Eq. (12). Finally the measured and calculated four output Stokes parameters are compared in the wavelength range from 1540 to 1560 nm with 1 nm step size in Fig. 2.

It is obvious that the calculated results are in excellent agreement with the measured ones. The maximum relative difference between calculated and measured Stokes parameters is less than 1.5%. Without doubt, Mueller matrix can be determined from Eq. (14) in high accuracy using these calculated data. Since the Mueller matrices have been measured, the PDL can be calculated using $\mathrm{PDL} = 10\log\frac{m_{11}+\sqrt{m_{12}^2+m_{13}^2+m_{14}^2}}{m_{11}-\sqrt{m_{12}^2+m_{13}^2+m_{14}^2}}$ [14], and the results are shown in Fig. 3.

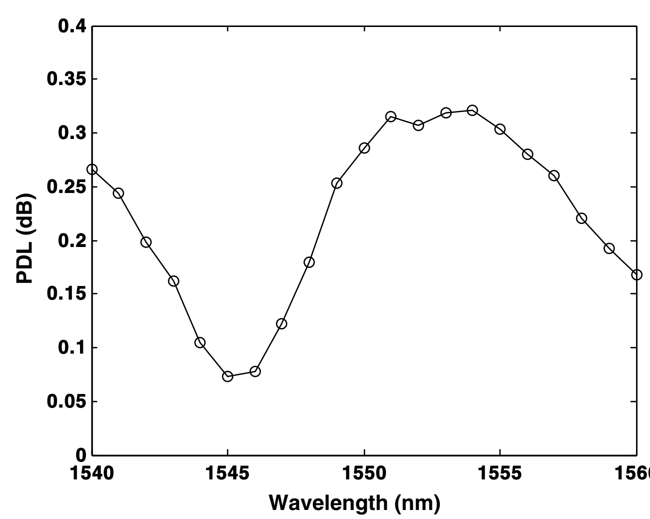

Fig. 3. Measured PDL evolution with respect to optical wavelength.

6. Conclusion

We studied the relationships between a set of input polarization states and their corresponding outputs based on the fact that Mueller matrix for an optical fiber system with both birefringence and PDL/G meets the Lorentz transformation. Based on these relationships, we can measure Mueller matrix using three arbitrary inputs, which are sufficient but not redundant. Because the measurement accuracy is affected by various errors in real tests, the requirements regarding three inputs for the optimum accuracy are theoretically demonstrated. Experimental results on an optical fiber system consisting of SMFs and an optical isolator confirm the validity of our theory.

Acknowledgements

This work is partially supported by the project 042 101 0015 of Agency for Science, Technology and Research (A*Star), Singapore, and Open Fund of Key Laboratory of Optical Communication and Lightwave Technologies, Beijing University of Posts and Telecommunications, Ministry of Education, PR China.

Appendix A

In this appendix, the solvable condition and the condition for the optimum accuracy of Eq. (12) will be discussed.

Eq. (13) gives the direct condition that Eq. (12) can be solved, which is that the determinant of the coefficient matrix cannot be equal to 0. And inversely we can consider the condition for the optimum accuracy is that the absolute value of this determinant can be up to its maximum. Because Eq. (13) only indicates the requirements of the output Stokes parameters, we have to find the requirement of the input Stokes parameters at first. From Eq. (12), we have

$$\begin{pmatrix} \mathrm{i}s_{\text{out}0} & s_{\text{out}1} & s_{\text{out}2} & s_{\text{out}3} \\ \mathrm{i}t_{\text{out}0} & t_{\text{out}1} & t_{\text{out}2} & t_{\text{out}3} \\ \mathrm{i}u_{\text{out}0} & u_{\text{out}1} & u_{\text{out}2} & u_{\text{out}3} \\ A_{\text{out}0} & A_{\text{out}1} & A_{\text{out}2} & A_{\text{out}3} \end{pmatrix} \begin{pmatrix} \mathrm{i}v_{\text{out}0} \\ v_{\text{out}1} \\ v_{\text{out}2} \\ v_{\text{out}3} \end{pmatrix} = \sqrt{|\mathbf{M}|} \begin{pmatrix} \mathrm{i}s_{\text{in}0} & s_{\text{in}1} & s_{\text{in}2} & s_{\text{in}3} \\ \mathrm{i}t_{\text{in}0} & t_{\text{in}1} & t_{\text{in}2} & t_{\text{in}3} \\ \mathrm{i}u_{\text{in}0} & u_{\text{in}1} & u_{\text{in}2} & u_{\text{in}3} \\ \sqrt{|\mathbf{M}|}A_{\text{in}0} & \sqrt{|\mathbf{M}|}A_{\text{in}1} & \sqrt{|\mathbf{M}|}A_{\text{in}2} & \sqrt{|\mathbf{M}|}A_{\text{in}3} \end{pmatrix} \begin{pmatrix} \mathrm{i}v_{\text{in}0} \\ v_{\text{in}1} \\ v_{\text{in}2} \\ v_{\text{in}3} \end{pmatrix}. \tag{A1}$$

Take the notations that

$$\mathbf{F}_{\text{out}} = \begin{pmatrix} \mathrm{i}s_{\text{out}0} & s_{\text{out}1} & s_{\text{out}2} & s_{\text{out}3} \\ \mathrm{i}t_{\text{out}0} & t_{\text{out}1} & t_{\text{out}2} & t_{\text{out}3} \\ \mathrm{i}u_{\text{out}0} & u_{\text{out}1} & u_{\text{out}2} & u_{\text{out}3} \\ A_{\text{out}0} & A_{\text{out}1} & A_{\text{out}2} & A_{\text{out}3} \end{pmatrix}, \quad \mathbf{F}_{\text{in}} = \begin{pmatrix} \mathrm{i}s_{\text{in}0} & s_{\text{in}1} & s_{\text{in}2} & s_{\text{in}3} \\ \mathrm{i}t_{\text{in}0} & t_{\text{in}1} & t_{\text{in}2} & t_{\text{in}3} \\ \mathrm{i}u_{\text{in}0} & u_{\text{in}1} & u_{\text{in}2} & u_{\text{in}3} \\ A_{\text{in}0} & A_{\text{in}1} & A_{\text{in}2} & A_{\text{in}3} \end{pmatrix} \text{ and } \widetilde{\mathbf{F}}_{\text{in}} = \begin{pmatrix} \mathrm{i}s_{\text{in}0} & s_{\text{in}1} & s_{\text{in}2} & s_{\text{in}3} \\ \mathrm{i}t_{\text{in}0} & t_{\text{in}1} & t_{\text{in}2} & t_{\text{in}3} \\ \mathrm{i}u_{\text{in}0} & u_{\text{in}1} & u_{\text{in}2} & u_{\text{in}3} \\ \sqrt{|\mathbf{M}|}A_{\text{in}0} & \sqrt{|\mathbf{M}|}A_{\text{in}1} & \sqrt{|\mathbf{M}|}A_{\text{in}2} & \sqrt{|\mathbf{M}|}A_{\text{in}3} \end{pmatrix},$$

then $\mathbf{F}_{\text{out}}\vec{V}_{\text{out}} = \sqrt{|\mathbf{M}|}\widetilde{\mathbf{F}}_{\text{in}}\vec{V}_{\text{in}}$. Due to $\vec{V}_{\text{out}} = \mathbf{M}\vec{V}_{\text{in}}$, we have

$$\mathbf{F}_{\text{out}}\mathbf{M}\vec{V}_{\text{in}} = \sqrt{|\mathbf{M}|}\widetilde{\mathbf{F}}_{\text{in}}\vec{V}_{\text{in}}. \tag{A2}$$

Because $\vec{V}_{\text{in}}$ can be arbitrary, the square matrices on the left and right sides must be equal

$$\mathbf{F}_{\text{out}}\mathbf{M} = \sqrt{|\mathbf{M}|}\widetilde{\mathbf{F}}_{\text{in}}. \tag{A3}$$

Finally we obtain a relation as

$$|\mathbf{F}_{\text{out}}| = |\mathbf{M}||\widetilde{\mathbf{F}}_{\text{in}}| = |\mathbf{M}|^{\frac{3}{2}}|\mathbf{F}_{\text{in}}|. \tag{A4}$$

Based on Eq. (A4), we can study the above-mentioned conditions using the inputs if we consider the matrix determinant as the criteria. By calculation, we can obtain

$$|\mathbf{F}_{\text{in}}| = s_{\text{in}0}^2 t_{\text{in}0}^2 u_{\text{in}0}^2 \{[\widehat{S}_{\text{in}} \cdot (\widehat{T}_{\text{in}} \times \widehat{U}_{\text{in}})]^2 - |(\widehat{T}_{\text{in}} - \widehat{S}_{\text{in}}) \times (\widehat{U}_{\text{in}} - \widehat{S}_{\text{in}})|^2\}. \tag{A5}$$

Here "^" indicates a normalized 3-D Stokes vector. In Fig. A1, we show the geometrical meaning of Eq. (A5). The first item inside the bracket of Eq. (A5) stands for the square of the volume of a parallelepiped composed of $\widehat{S}_{\text{in}}$, $\widehat{T}_{\text{in}}$ and $\widehat{U}_{\text{in}}$; the second item denotes the 4-time value of the square of the area of the triangle shadowed in the figure.

A.1. The solvable condition

The solvable condition of Eq. (12) is $|\mathbf{F}_{\text{out}}| \neq 0$, viz., $|\mathbf{F}_{\text{in}}| \neq 0$ which can be observed from Eq. (A4). From Eq. (A5), we know there are two cases that $|\mathbf{F}_{\text{in}}|$ is possible to be equal to 0, which are discussed as below.

(1) $\widehat{S}_{\text{in}} \cdot (\widehat{T}_{\text{in}} \times \widehat{U}_{\text{in}}) = 0$ and $(\widehat{T}_{\text{in}} - \widehat{S}_{\text{in}}) \times (\widehat{U}_{\text{in}} - \widehat{S}_{\text{in}}) = 0$

From the first equation, we know $\widehat{S}_{\text{in}}$, $\widehat{T}_{\text{in}}$ and $\widehat{U}_{\text{in}}$ are coplanar in Stokes space. In this case, we know the second equation stands for the shadow area in Fig. A2. Then

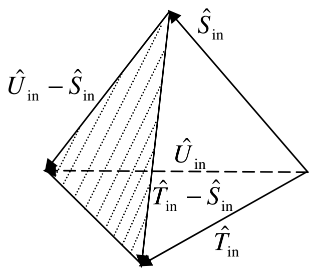

Fig. A1. Geometrical meaning of Eq. (A5).

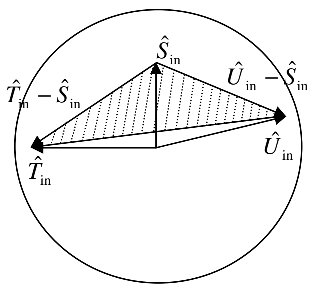

Fig. A2. Geometrical meaning of Eq. (A6).

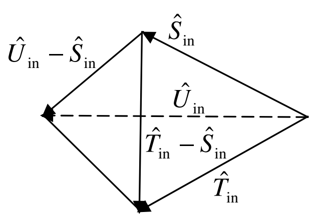

Fig. A3. Geometrical relations under condition 2.

$$|\mathbf{F}_{in}| = -4s_{in0}^2 t_{in0}^2 u_{in0}^2 (\text{Area of the shadow region})^2. \quad \text{(A6)}$$

So in this case, when the area of the shadow region is zero, Eq. (12) can not be solved. This will happen when $\widehat{T}_{in} - \widehat{S}_{in}$ is parallel to $\widehat{U}_{in} - \widehat{S}_{in}$. While the inputs are completely polarized lights, which are typically used in real measurements, the condition becomes the three inputs can not be linearly superposed.

(2) $|\widehat{S}_{in} \cdot (\widehat{T}_{in} \times \widehat{U}_{in})| = |(\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in})|$

We can easily derive $\widehat{S}_{in} \cdot \lfloor (\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in}) \rfloor = \widehat{S}_{in} \cdot (\widehat{T}_{in} \times \widehat{U}_{in})$. This means two relations must be satisfied as: $|\widehat{S}_{in}| = 1$ and $\widehat{S}_{in}$ is parallel to $(\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in})$ (Viz., $\widehat{S}_{in}$ is orthogonal to $\widehat{T}_{in} - \widehat{S}_{in}$ and $\widehat{U}_{in} - \widehat{S}_{in}$). In Fig. A3, it is easily observed that $|\widehat{T}_{in}|$ and $|\widehat{U}|_{in}$ will larger than 1 under such relations. But this is impossible. So $|\mathbf{F}_{in}|$ will not be 0 in this case.

In conclusion, the solvable condition for Eq. (12) is the three inputs are not linearly superposed in Stokes space while they are completely polarized.

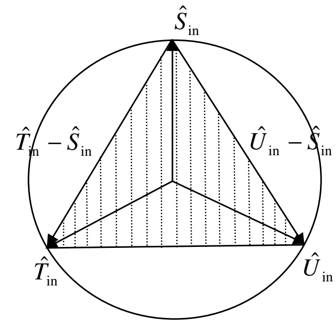

Fig. A4. Geometrical relations of three normalized inputs for optimum accuracy.

A.2. The condition for optimum accuracy

We consider the condition for optimum accuracy is that the absolute value of the determinant achieves its maximum. Based on Eq. (A4), abs($|\mathbf{F}_{out}|$) and abs($|\mathbf{F}_{in}|$) reach their maximum simultaneously.

From Eq. (A5), we easily know if one item inside the bracket is equal to 0, and another item is up to its maximum value at the same time, abs($|\mathbf{F}_{in}|$) reaches its maximum.

(1) $\widehat{S}_{in} \cdot (\widehat{T}_{in} \times \widehat{U}_{in}) = 0$

Now $\widehat{S}_{in}$, $\widehat{T}_{in}$ and $\widehat{U}_{in}$ are coplanar, and then

$$\text{abs}(|\mathbf{F}_{in}|) = s_{in0}^2 t_{in0}^2 u_{in0}^2 |(\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in})|^2. \quad \text{(A7)}$$

By calculation, we easily confirm $|(\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in})|^2$ will achieve its maximum value in this case. The requirements are that $\widehat{S}_{in}$, $\widehat{T}_{in}$ and $\widehat{U}_{in}$ are with 120 degrees between them, and their magnitudes are equal to 1, as shown in the Fig. A4.

At present, we easily obtain that abs($|\mathbf{F}_{in}|$) = $6\frac{3}{4} s_{in0}^2 t_{in0}^2 u_{in0}^2$. As a comparison, abs($|\mathbf{F}_{in}|$) = $4 s_{in0}^2 t_{in0}^2 u_{in0}^2$ when three inputs are $(1\,0\,0)^T$, $(-1\,0\,0)^T$ and $(0\,1\,0)^T$.

(2) $(\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in}) = 0$

In this case, we already know $\widehat{S}_{in} \cdot \lfloor (\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in}) \rfloor = \widehat{S}_{in} \cdot (\widehat{T}_{in} \times \widehat{U}_{in})$. If $(\widehat{T}_{in} - \widehat{S}_{in}) \times (\widehat{U}_{in} - \widehat{S}_{in}) = 0$, then $\widehat{S}_{in} \cdot (\widehat{T}_{in} \times \widehat{U}_{in}) = 0$. So abs($|\mathbf{F}_{in}|$) $= s_{in0}^2 t_{in0}^2 u_{in0}^2 \text{abs}\{[\widehat{S}_{in} \cdot (\widehat{T}_{in} \times \widehat{U}_{in})]^2\} = 0$.

In conclusion, the maximum value will be achieved when $\widehat{S}_{in}$, $\widehat{T}_{in}$ and $\widehat{U}_{in}$ are coplanar, their angles between each other are 120° and their magnitudes are equal to 1.

References

[1] N. Gisin, B. Huttner, Opt. Commun. 142 (1997) 119.
[2] B. Huttner, C. Geiser, N. Gisin, IEEE J. Sel. Top. Quantum Electron. 6 (2000) 317.
[3] A. Gerrard, J.M. Burch, Introduction to Matrix Methods in Optics, Wiley, London, 1975.
[4] R.M. Jopson, L.E. Nelson, H. Kogelnik, IEEE Photon. Technol. Lett. 11 (1999) 1153.

[5] Shih-Yau Lu, Russell A. Chipman, J. Opt. Soc. Am. A 13 (1996) 1106.
[6] R. Simon, Opt. Commun. 77 (1991) 349.
[7] R. Clark Jones, J. Opt. Soc. Am 37 (1947) 110.
[8] H. Dong, P. Shum, M. Yan, J.Q. Zhou, G.X. Ning, Y.D. Gong, C.Q. Wu, Opt. Express 14 (2006) 5067. http://www.opticsinfobase.org/abstract.cfm?URI=oe-14-12-5067.
[9] Richard Barakat, J. Opt. Soc. Am. 53 (1963) 317.
[10] H. Dong, P. Shum, M. Yan, G. Ning, Y. Gong, C. Wu, Opt. Express 13 (2005) 8875. http://www.opticsexpress.org/abstract.cfm?URI=OPEX-13-22-8875.
[11] G. Stephenson, C.W. Kilmister, Special Relativity for Physicists, Longmans, Green and Co., 1960.
[12] Richard Barakat, Opt. Commun. 38 (1981) 159.
[13] Barton J. Howell, Appl. Opt. 18 (1979) 809.
[14] B.L. Heffner, IEEE Photon. Technol. Lett. 4 (1992) 451.