Vector theory of stimulated Raman scattering and its application to fiber-based Raman amplifiers

Qiang Lin and Govind P. Agrawal

Institute of Optics, University of Rochester, Rochester, New York 14627

Received December 4, 2002; revised manuscript received April 3, 2003

A vector theory of the stimulated Raman scattering process is developed for describing the polarization effects in fiber-based Raman amplifiers. We use this theory to show that polarization-mode dispersion (PMD) induces large fluctuations in an amplified signal. It is found that PMD-induced fluctuations follow a log-normal distribution. We also discuss the random nature of the polarization-dependent gain (PDG) in Raman amplifiers. Using the concept of a PDG vector, we find the probability distribution of PDG in an analytic form and use it to show that both the mean and the standard deviation of PDG depend on the PMD parameter inversely when the effective fiber length is much larger than the PMD diffusion length. We apply our theory to study how PDG can be reduced by scrambling pump polarization randomly and show that the mean value of PDG is directly proportional to the degree of pump polarization.

OCIS codes: 060.4370, 060.2330, 190.5650, 000.5490.

1. INTRODUCTION

The Raman effect, first observed in 1928,[1] has attracted considerable attention since 1962 when the phenomenon of stimulated Raman scattering (SRS) was discovered.[2] SRS was first observed in silica fibers in 1972 (Ref. 3), and by 1981 it was used to make fiber-based Raman amplifiers capable of providing more than 30-dB gain.[4] Such Raman amplifiers have attracted considerable attention recently[5–7] because of their potential for providing a relatively flat gain over a wide bandwidth. The theoretical treatment of Raman amplifiers is often based on a scalar approach[6] even though the Raman gain is known to be polarization dependent.[8–13] A scalar approach can be justified if the polarization states of the pump and the signal fields do not change along the fiber. This is, however, not the case in most fibers in which birefringence fluctuations lead to randomization of the state of polarization (SOP) of any optical field. This effect is known as polarization-mode dispersion (PMD) and has been studied extensively in recent years.[14–17] Although the effects of PMD on Raman amplification have been observed experimentally,[10–13] a vector theory of the SRS process has not yet been fully developed.

In this paper we develop a vector theory of Raman amplification capable of including the PMD-induced random evolution of the pump and signal polarization states.[18,19] We use this theory to show that the amplified signal fluctuates over a wide range because of PMD, and the average gain is significantly lower than that expected in the absence of PMD. Based on this theory, we find the statistics of polarization-dependent gain (PDG) and its relationship with the operating parameters of Raman amplifiers. The paper is organized as follows. In Section 2 we develop the basic theory using the Stokes-vector formalism and discuss the simplifying approximations made for obtaining the analytical results. The average value of the amplified signal and the variance of its PMD-induced fluctuations are discussed in Section 3. The probability density of signal fluctuations is shown to be a log-normal distribution in Section 4. In Section 5 we consider the statistics of PDG and show that both the average and rms values of PDG can be found in an analytic form. We apply in Section 6 the vector theory to the case in which pump polarization is scrambled randomly to reduce the influence of PMD effects. The main results are summarized in Section 7.

2. GENERAL VECTOR THEORY

In the field of nonlinear optics, the polarization induced in a dielectric medium is expanded in powers of the optical field $\mathbf{E}(\mathbf{r}, t)$ as

$$\mathbf{P}(\mathbf{r}, t) = \mathbf{P}^{(1)}(\mathbf{r}, t) + \mathbf{P}^{(2)}(\mathbf{r}, t) + \mathbf{P}^{(3)}(\mathbf{r}, t) + \cdots, \tag{2.1}$$

where $\mathbf{P}^{(1)}$ represents the linear contribution and the other terms account for the second-, third-, and higher-order nonlinear effects. We assume that the medium exhibits inversion symmetry so that the second-order nonlinear effects do not occur, and $\mathbf{P}^{(2)} = 0$. The third-order nonlinear polarization in a medium such as silica glass can be written in its most general form as[20,21]

$$\begin{aligned}\mathbf{P}^{(3)}(\mathbf{r}, t) &= \frac{\varepsilon_0}{2}\sigma[\mathbf{E}(\mathbf{r}, t)\cdot\mathbf{E}(\mathbf{r}, t)]\mathbf{E}(\mathbf{r}, t) \\ &+ \mathbf{E}(\mathbf{r}, t)\int_0^\infty \varepsilon_0 a(\tau)[\mathbf{E}(\mathbf{r}, t-\tau)\cdot\mathbf{E}(\mathbf{r}, t-\tau)]\mathrm{d}\tau \\ &+ \mathbf{E}(\mathbf{r}, t)\cdot\int_0^\infty \varepsilon_0 b(\tau)\mathbf{E}(\mathbf{r}, t-\tau)\mathbf{E}(\mathbf{r}, t \\ &- \tau)\mathrm{d}\tau, \end{aligned} \tag{2.2}$$

where $a(\tau)$ and $b(\tau)$ govern the delayed Raman response (related to nuclear motion) whereas σ accounts for the instantaneous electronic response of the nonlinear medium.

In a Raman amplifier, the pump and signal waves propagate simultaneously, and the total field is given by

$$\mathbf{E} = \mathrm{Re}[\mathbf{E}_p \exp(-i\omega_p t) + \mathbf{E}_s \exp(-i\omega_s t)], \quad (2.3)$$

where $\mathbf{E}_p$ and $\mathbf{E}_s$ are the slowly varying envelopes for the pump and signal fields oscillating at frequencies ω_p and ω_s, respectively. Writing $\mathbf{P}^{(3)}$ also in the same form as

$$\mathbf{P}^{(3)} = \mathrm{Re}[\mathbf{P}_p \exp(-i\omega_p t) + \mathbf{P}_s \exp(-i\omega_s t)], \quad (2.4)$$

the nonlinear polarization at the pump and signal frequencies is found to be

$$\begin{aligned}\mathbf{P}_j(\omega_j) = {} & \frac{\varepsilon_0}{8}[\sigma + 2\tilde{b}(0)](\mathbf{E}_j \cdot \mathbf{E}_j)\mathbf{E}_j^* \\ & + \frac{\varepsilon_0}{4}[\sigma + 2\tilde{a}(0) + \tilde{b}(0)](\mathbf{E}_j^* \cdot \mathbf{E}_j)\mathbf{E}_j \\ & + \frac{\varepsilon_0}{4}[\sigma + 2\tilde{a}(0) + \tilde{b}(\omega_j - \omega_m)](\mathbf{E}_m^* \cdot \mathbf{E}_m)\mathbf{E}_j \\ & + \frac{\varepsilon_0}{4}[\sigma + \tilde{b}(0) + \tilde{b}(\omega_j - \omega_m)](\mathbf{E}_m \cdot \mathbf{E}_j)\mathbf{E}_m^* \\ & + \frac{\varepsilon_0}{4}[\sigma + 2\tilde{a}(\omega_j - \omega_m) + \tilde{b}(0)] \\ & \times (\mathbf{E}_m^* \cdot \mathbf{E}_j)\mathbf{E}_m, \end{aligned} \quad (2.5)$$

where $j, m = p$ or $s (j \neq m)$ and

$$\tilde{a}(\omega) = \int_0^\infty a(\tau)\exp(i\omega\tau)\mathrm{d}\tau,$$

$$\tilde{b}(\omega) = \int_0^\infty b(\tau)\exp(i\omega\tau)\mathrm{d}\tau \quad (2.6)$$

are the Fourier transforms of the Raman response functions $a(\tau)$ and $b(\tau)$, respectively. In obtaining Eq. (2.5), the terms containing $\tilde{a}(2\omega_p)$, $\tilde{a}(\omega_p + \omega_s)$, $\tilde{b}(2\omega_p)$, and $\tilde{b}(\omega_p + \omega_s)$ were neglected because of their relatively small magnitudes.

We need to make several simplifying assumptions before the vector theory of Raman amplification can be used successfully in practice. Both the pump and signal fields are generally time dependent in practice. We assume that they vary with time on a time scale long enough that the effects of group-velocity dispersion are negligible. The shortest time scale is the bit duration T_B related inversely to bit rate B of a light-wave system. Dispersive effects are negligible when fiber length L is a small fraction of the dispersion length $L_D = T_B^2/|\beta_2|$, where $|\beta_2|$ is the group-velocity dispersion parameter. Even at $B = 40$ Gbit/s, $L_D > 100$ km if we use $|\beta_2| = 5$ ps^2/km, and the effects of group-velocity dispersion can be neglected for $L < 20$ km.

If we use Eqs. (2.1), (2.3), and (2.4) in the Maxwell equations together with $\mathbf{P}^{(2)} = 0$, we find that $\mathbf{E}_p$ and $\mathbf{E}_s$ satisfy the nonlinear Helmholtz equations

$$\nabla^2\mathbf{E}_j + \frac{\omega_j^2}{c^2}\vec{\varepsilon}_j\mathbf{E}_j = -\frac{\omega_j^2}{\varepsilon_0 c^2}\mathbf{P}_j, \quad (2.7)$$

where ε_0 is the vacuum permittivity and $\vec{\varepsilon}_j$ is the linear part of the dielectric constant resulting from $\mathbf{P}^{(1)}$ in Eq. (2.1). Note that its tensorial nature is important to account for the PMD effects that have their origin in the birefringence of silica fibers.

Both $\mathbf{E}_p$ and $\mathbf{E}_s$ vectors evolve along the fiber length. Their magnitudes change because of SRS whereas their SOPs change because of birefringence. It is common to choose the z axis along the fiber length and assume that $\mathbf{E}_p$ and $\mathbf{E}_s$ lie in the x–y plane. This assumption amounts to a neglect of the longitudinal component of the two vectors and is justified in practice as long as the spatial size of the fiber mode is larger than the optical wavelength. In the Jones-matrix notation of Ref. 14, the pump and signal fields at any point $\mathbf{r}$ inside the fiber can be written as

$$\mathbf{E}_p(\mathbf{r}) = F_p(x, y)|A_p\rangle\exp(ik_p z),$$

$$\mathbf{E}_s(\mathbf{r}) = F_s(x, y)|A_s\rangle\exp(ik_s z), \quad (2.8)$$

where $F_p(x, y)$ and $F_s(x, y)$ represent the fiber-mode profile, k_p and k_s are propagation constants, and the Jones vectors $|A_p\rangle$ and $|A_s\rangle$ are two-dimensional column vectors representing the two components of the electric field in the x–y plane. Since F_p and F_s do not change with z, we need to consider only the evolution of $|A_p\rangle$ and $|A_s\rangle$ along the fiber.

We substitute Eqs. (2.8) back into Eq. (2.7), integrate over the transverse mode distribution in the x–y plane, and assume $|A_p\rangle$ and $|A_s\rangle$ to be slowly varying functions of z so that we can neglect their second-order derivative with respect to z. We write $\vec{\varepsilon}$ on the basis of Pauli matrices as[14]

$$(\omega_j^2/c^2)\vec{\varepsilon}_j = (k_j + i\alpha_j/2)^2\sigma_0 - k_j\omega_j\boldsymbol{\beta}\cdot\boldsymbol{\sigma}, \quad (2.9)$$

where σ_0 is a unit matrix. The vector $\boldsymbol{\sigma}$ is formed by use of the Pauli matrices as $\boldsymbol{\sigma} = \sigma_1\hat{e}_1 + \sigma_2\hat{e}_2 + \sigma_3\hat{e}_3$, where $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ are the three unit vectors in the Stokes space and

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (2.10)$$

In the Jones-matrix notation, we obtain the following vector equations governing the evolution of the pump and signal fields inside an optical fiber:

$$\xi\frac{\mathrm{d}|A_j\rangle}{\mathrm{d}z} = -\frac{\alpha_j}{2}|A_j\rangle - \frac{i}{2}\omega_j\boldsymbol{\beta}\cdot\boldsymbol{\sigma}|A_j\rangle + \frac{i\gamma_{jj}}{3}\left[2\langle A_j|A_j\rangle + \frac{\kappa_b}{\kappa_a}|A_j^*\rangle\langle A_j^*|\right]|A_j\rangle + \frac{2i\gamma_{jm}}{3}[(1+\delta_b)\langle A_m|A_m\rangle + (1+\delta_a)|A_m\rangle\langle A_m| + (\kappa_b/\kappa_a+\delta_b)|A_m^*\rangle\langle A_m^*|]|A_j\rangle + \frac{\zeta}{2}[g_2(\langle A_m|A_m\rangle + |A_m^*\rangle\langle A_m^*|) + g_1|A_m\rangle\langle A_m|]|A_j\rangle, \quad (2.11)$$

where $j, m = p$ or s $(m \neq j)$, α_j accounts for fiber losses, and $\zeta = 1$ when $j = s$ but $\zeta = -\omega_p/\omega_s$ when $j = p$ because of the SRS process. The vectors $\langle A|$ and $|A^*\rangle$ are Hermitian and complex conjugates of $|A\rangle$, respectively. The PMD effects in Eq. (2.11) are governed by the birefringence vector $\boldsymbol{\beta}$.[22] $\xi = \pm 1$ depending on the pumping configuration. In the following analysis we assume the signal propagates forward but the pump can propagate forward ($\xi = 1$) or backward ($\xi = -1$), depending on the pumping configuration.

The nonlinear effects in Eq. (2.11) are governed by the parameters

$$\kappa_a = \sigma + 2\tilde{a}(0) + \tilde{b}(0), \quad \kappa_b = \sigma + 2\tilde{b}(0), \quad (2.12)$$

$$\delta_a = 2\{\mathrm{Re}[\tilde{a}(\Omega_R)] - \tilde{a}(0)\}/\kappa_a,$$

$$\delta_b = \{\mathrm{Re}[\tilde{b}(\Omega_R)] - \tilde{b}(0)\}/\kappa_a, \quad (2.13)$$

where $\Omega_R = \omega_p - \omega_s$ is the Raman shift. These parameters are responsible for self-phase modulation and cross-phase modulation (XPM) and lead to the phenomenon of nonlinear polarization rotation (NPR). Usually the instantaneous Kerr response dominates and σ is so large that $\kappa_a \approx \kappa_b$.[23] We have also introduced

$$\gamma_{jm} = 3\omega_j^2\omega_m\kappa_a/(8c^4k_jk_mA_{\mathrm{eff}}) \quad (2.14)$$

as the self-phase modulation ($m = j$) and XPM ($m \neq j$) nonlinear parameters at the pump ($j = p$) and signal ($j = s$) frequencies, where A_{eff} is the effective core area of the optical fiber assumed to be the same for both the pump and the signal.[6]

The Raman-gain parameters g_1 and g_2 in Eq. (2.11) account for the contributions from the isotropic and anisotropic nuclear response, respectively, and are defined as[20]

$$g_1 = \omega_s^2\omega_p\,\mathrm{Im}[\tilde{a}(\Omega_R)]/(c^4k_pk_sA_{\mathrm{eff}}),$$

$$g_2 = \omega_s^2\omega_p\,\mathrm{Im}[\tilde{b}(\Omega_R)]/(2c^4k_pk_sA_{\mathrm{eff}}). \quad (2.15)$$

In general, g_2 is much smaller than g_1 for optical fibers.[9,21]

Equation (2.11) looks complicated in the Jones-matrix formalism. It can be simplified considerably by writing it in the Stokes space.[14] After introducing the Stokes vectors for the pump and signal as

$$\mathbf{P} = \langle A_p|\boldsymbol{\sigma}|A_p\rangle \equiv P_1\hat{e}_1 + P_2\hat{e}_2 + P_3\hat{e}_3, \quad (2.16)$$

$$\mathbf{S} = \langle A_s|\boldsymbol{\sigma}|A_s\rangle \equiv S_1\hat{e}_1 + S_2\hat{e}_2 + S_3\hat{e}_3, \quad (2.17)$$

and using the relations

$$|A^*\rangle\langle A^*| = |A\rangle\langle A| - \langle A|\sigma_3|A\rangle\sigma_3, \quad (2.18)$$

$$|A\rangle\langle A| = \{\langle A|A\rangle + \langle A|\boldsymbol{\sigma}|A\rangle\cdot\boldsymbol{\sigma}\}/2, \quad (2.19)$$

we obtain the following two vector equations governing the dynamics of $\mathbf{P}$ and $\mathbf{S}$ in the Stokes space:

$$\xi\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}z} = -\alpha_p\mathbf{P} - \frac{\omega_p}{2\omega_s}g_1[(1+3\mu)S_0\mathbf{P} + (1+\mu)P_0\mathbf{S} - 2\mu P_0\mathbf{S}_3] + (\omega_p\boldsymbol{\beta} + \mathbf{W}_p)\times\mathbf{P}, \quad (2.20)$$

$$\frac{\mathrm{d}\mathbf{S}}{\mathrm{d}z} = -\alpha_s\mathbf{S} + \frac{g_1}{2}[(1+3\mu)P_0\mathbf{S} + (1+\mu)S_0\mathbf{P} - 2\mu S_0\mathbf{P}_3] + (\omega_s\boldsymbol{\beta} + \mathbf{W}_s)\times\mathbf{S}, \quad (2.21)$$

where $P_0 = |\mathbf{P}|$ and $S_0 = |\mathbf{S}|$ are the pump and signal powers, $\mathbf{P}_3 = P_3\hat{e}_3$, $\mathbf{S}_3 = S_3\hat{e}_3$ represent projections along $\hat{e}_3$, and

$$\mathbf{W}_p = \frac{2}{3}[\gamma_{pp}\mathbf{P}_3 + 2\gamma_{ps}(1+\delta_b)\mathbf{S}_3 - \gamma_{ps}(2+\delta_a+\delta_b)\mathbf{S}], \quad (2.22)$$

$$\mathbf{W}_s = \frac{2}{3}[\gamma_{ss}\mathbf{S}_3 + 2\gamma_{sp}(1+\delta_b)\mathbf{P}_3 - \gamma_{sp}(2+\delta_a+\delta_b)\mathbf{P}]. \quad (2.23)$$

The vectors $\mathbf{W}_p$ and $\mathbf{W}_s$ account for the self-phase modulation-induced and XPM-induced NPR. The parameter $\mu = g_2/g_1$ is a measure of the anisotropy and depends in general on the pump-signal detuning Ω_R. It has a value of approximately 0.012 for silica fibers at the Raman gain peak.[20] In most of the previous work μ has been set to zero.[8,18]

Because of fiber birefringence, both $\mathbf{P}$ and $\mathbf{S}$ rotate on the Poincaré sphere around the same axis but with different rates as dictated by the magnitude of $\omega_p\boldsymbol{\beta}$ and $\omega_s\boldsymbol{\beta}$. However, the SRS process depends only on the relative orientation of $\mathbf{P}$ and $\mathbf{S}$. To simplify the following analysis, we chose to work in a rotating frame in which the pump Stokes vector $\mathbf{P}$ is not affected by birefringence. In the Stokes space, the required transformation is $\mathbf{V} = \vec{R}\mathbf{V}'$, where $\mathbf{V}$ is an arbitrary vector in the Stokes space and the rotation matrix $\vec{R}$ evolves with z as

$$\frac{\mathrm{d}\vec{R}}{\mathrm{d}z} = \xi\omega_p\boldsymbol{\beta}\times\vec{R}. \quad (2.24)$$

In practice, the beat length of residual birefringence (~1 m) and its correlation length (~10 m) are much smaller than the beat length of NPR (~10 km depending on optical powers). As a result, rotations of Stokes vectors induced by fiber birefringence are so fast compared with that induced by NPR that one can average over them.

Appendix A provides details of the averaging procedure. The averaged equations become

$$\xi \frac{d\mathbf{P}}{dz} = -\alpha_p \mathbf{P} - \frac{\omega_p}{2\omega_s} g_1[(1 + 3\mu)S_0\mathbf{P} + (1 + \mu/3)P_0\mathbf{S}] - \varepsilon_{ps}\mathbf{S} \times \mathbf{P}, \quad (2.25)$$

$$\frac{d\mathbf{S}}{dz} = -\alpha_s \mathbf{S} + \frac{g_1}{2}[(1 + 3\mu)P_0\mathbf{S} + (1 + \mu/3)S_0\mathbf{P}] - (\Omega_R \mathbf{B} + \varepsilon_{sp}\mathbf{P}) \times \mathbf{S}, \quad (2.26)$$

where $\varepsilon_{jm} = 2\gamma_{jm}(4 + 3\delta_a + \delta_b)/9$, $\Omega_R = \xi\omega_p - \omega_s$, and $\mathbf{B}$ is related to $\boldsymbol{\beta}$ through $\tilde{R}$.

Equations (2.25) and (2.26) describe SRS under quite general conditions. We make two further simplifications in the following analysis. We neglect both the pump depletion and the signal-induced XPM on the pump because the pump power is much larger than the signal power in practice. The pump equation (2.25) then contains only the loss term and can be easily integrated. The effect of fiber losses is to reduce the magnitude of $\mathbf{P}$ but the direction of $\mathbf{P}$ remains fixed in the rotating frame.

Equation (2.26) shows clearly that the Raman gain is polarization dependent. The gain coefficient varies from $g_1(1 + 5\mu/3)$ to $4\mu g_1/3$, depending on the angle between the Stokes vectors of the pump and the signal. Random variations in the fiber birefringence change the relative orientation between $\mathbf{S}$ and $\mathbf{P}$ and produce random changes in the Raman gain. However, the last term, $\varepsilon_{sp}\mathbf{P} \times \mathbf{S}$, in Eq. (2.26) accounts for the XPM-induced NPR and does not affect the Raman gain because of its deterministic nature. We can eliminate this term by making a further transformation

$$\mathbf{V} = \exp\left\{-\varepsilon_{sp}\left[\int_0^z P_0(z)dz\right]\hat{p} \times \right\}\mathbf{V}', \quad (2.27)$$

where $\hat{p}$ represents the unit vector on the Poincaré sphere in the direction of $\mathbf{P}$ and $\mathbf{V}$ is an arbitrary vector in the Stokes space. After doing so, Eq. (2.26) reduces to

$$\frac{d\mathbf{S}}{dz} = -\alpha_s \mathbf{S} + \frac{g_1}{2}[(1 + 3\mu)P_0\mathbf{S} + (1 + \mu/3)S_0\mathbf{P}] - \Omega_R \mathbf{b} \times \mathbf{S}, \quad (2.28)$$

where $\mathbf{b}$ is related to $\mathbf{B}$ in Eq. (2.26) by a deterministic rotation. As optical fibers used for Raman amplification are much longer than the birefringence correlation length, $\mathbf{b}(z)$ can be modeled as a three-dimensional stochastic process whose first-order and second-order moments are given by

$$\langle \mathbf{b}(z) \rangle = 0, \qquad \langle \mathbf{b}(z_1)\mathbf{b}(z_2) \rangle = \frac{1}{3} D_p^2 \vec{I}\delta(z_2 - z_1), \quad (2.29)$$

where the angle bracket denotes an ensemble average, $\vec{I}$ is the second-order unit tensor, and D_p is the PMD parameter of the fiber. As discussed in Appendix B, we should treat all stochastic differential equations in the Stratonovich sense.[24]

Equation (2.28) can be further simplified by noting that the first two terms on its right-hand side do not change the direction of $\mathbf{S}$ and can be removed by a suitable transformation. Making the final transformation as

$$\mathbf{S} = \mathbf{s} \exp\left\{\int_0^z \left[\frac{g_1}{2}(1 + 3\mu)P_0(z) - \alpha_s\right]dz\right\}, \quad (2.30)$$

the dynamic equations governing the power and the SOP of the signal are given by

$$\frac{ds_0}{dz} = \frac{g_R}{2} P_0(z)\hat{p} \cdot \mathbf{s}, \quad (2.31)$$

$$\frac{d\mathbf{s}}{dz} = \frac{g_R}{2} P_0(z)s_0\hat{p} - \Omega_R \mathbf{b} \times \mathbf{s}, \quad (2.32)$$

where $g_R \equiv g_1(1 + \mu/3)$, $s_0 = |\mathbf{s}|$, and $\hat{p}$ is the input SOP of the pump.

Equations (2.31) and (2.32) apply for both the forward and the backward pumping schemes, but the z dependence of $P_0(z)$ and the magnitude of Ω_R depend on the pumping configuration. More specifically, $P_0(z) = P_{\text{in}} \exp(-\alpha_p z)$ and $\Omega_R = \omega_p - \omega_s$ in the case of forward pumping but $P_0(z) = P_{\text{in}} \exp[-\alpha_p(L - z)]$ and $\Omega_R = -(\omega_p + \omega_s)$ in the case of backward pumping, where P_{in} is the input pump power. In the absence of birefringence ($\mathbf{b}=0$), $\mathbf{s}$ remains oriented along $\hat{p}$, and we recover the scalar case.

3. AVERAGE RAMAN GAIN AND OUTPUT SIGNAL FLUCTUATIONS

Equations (2.31) and (2.32) can be used to calculate the power S_0 of the amplified signal as well as its SOP at any distance within the amplifier. When the birefringence vector $\mathbf{b}$ is z dependent, the solution depends on how $\mathbf{b}(z)$ changes. In the case of PMD, $\mathbf{b}(z)$ fluctuates with time. As a result, the amplified signal $S_0(L)$ at the output of an amplifier of length L also fluctuates. Such fluctuations would affect the performance of any light-wave system making use of Raman amplification. In this section we calculate the average and the variance of such PMD-induced fluctuations. We focus on the forward-pumping case for definiteness. All the results can be converted to the case of backward pumping by replacing α_p with $-\alpha_p$, P_{in} with $P_{\text{in}} \exp(-\alpha_p L)$, and $\Omega_R = \omega_p - \omega_s$ with $\Omega_R = -(\omega_p + \omega_s)$.

It is useful to introduce the instantaneous amplifier gain G defined as $G = S_0(L)/S_0(0)$. We use the vector theory to find the average Raman gain G_{av} and the signal variance σ_s^2 by using the definitions

$$G_{\text{av}} = \frac{\langle S_0(L) \rangle}{S_0(0)}, \qquad \sigma_s^2 = \frac{\langle S_0^2(L) \rangle}{\langle S_0(L) \rangle^2} - 1. \quad (3.1)$$

To calculate the average signal power $\langle S_0(L) \rangle$ at the end of a Raman amplifier of length L, we need to average Eqs. (2.31) and (2.32) using a well-known technique discussed in Ref. 24. Appendix B provides the details of the averaging procedure. The final result leads to the following two coupled but deterministic equations:

$$\frac{d\langle s_0 \rangle}{dz} = \frac{g_R}{2} P_0(z)\langle s_0 \cos\theta \rangle, \quad (3.2)$$

$$\frac{d\langle s_0 \cos\theta\rangle}{dz} = \frac{g_R}{2} P_0(z)\langle s_0\rangle - \eta\langle s_0 \cos\theta\rangle, \quad (3.3)$$

where $\eta = 1/L_d = D_p^2\Omega_R^2/3$, L_d is the PMD diffusion length, and θ is the angle between **P** and **S**.

Equations (3.2) and (3.3) are two linear first-order differential equations that can be easily integrated. When PMD effects are quite large, diffusion length L_d becomes so small that $\langle S_0 \cos\theta\rangle$ reduces to zero over a short fiber length of $\sim L_d$. The average gain is then given by (in decibels)

$$G_{av} = a[g_1(1 + 3\mu)P_{in}L_{eff}/2 - \alpha_s L], \quad (3.4)$$

where $a = 10/\ln 10 \approx 4.343$ and the effective amplifier length $L_{eff} = [1 - \exp(-\alpha_p L)]/\alpha_p$ is $< L$ because of pump losses. In this case, the PMD reduces the Raman gain coefficient to $g_1(1 + 3\mu)/2$, exactly the average of the copolarized $[g_1(1 + 5\mu/3)]$ and orthogonally polarized $(4\mu g_1/3)$ Raman gain coefficients.[8] If pump losses can be neglected ($\alpha_p = 0$), Eqs. (3.2) and (3.3) can be integrated analytically because P_0 becomes z independent. The average gain in this special case is given by

$$G_{av} = [\cosh(\kappa L/2) + \sinh(\kappa L/2)(g_R P_{in}\cos\theta_0 + \eta)/\kappa] \times \exp\{[g_1(1 + 3\mu)P_{in} - \eta - 2\alpha_s]L/2\}, \quad (3.5)$$

where θ_0 is the initial angle between **S** and **P** and $\kappa = [(g_R P_{in})^2 + \eta^2]^{1/2}$.

The variance of signal fluctuations requires the second-order moment $\langle S_0^2(L)\rangle$ of the amplified signal. Following the averaging procedure discussed in Appendix B, Eqs. (2.31) and (2.32) lead to the following set of three linear equations[24]:

$$\frac{d\langle s_0^2\rangle}{dz} = g_R P_0(z)\langle s_0^2 \cos\theta\rangle, \quad (3.6)$$

$$\frac{d\langle s_0^2 \cos\theta\rangle}{dz} = -\eta\langle s_0^2 \cos\theta\rangle + \frac{g_R}{2} P_0(z)[\langle s_0^2\rangle + \langle s_0^2 \cos^2\theta\rangle], \quad (3.7)$$

$$\frac{d\langle s_0^2 \cos^2\theta\rangle}{dz} = -3\eta\langle s_0^2 \cos^2\theta\rangle + \eta\langle s_0^2\rangle + g_R P_0(z)\langle s_0^2 \cos\theta\rangle. \quad (3.8)$$

These equations show that signal fluctuations have their origin in fluctuations of angle θ between the pump and the signal's Stokes vectors. The SRS process amplifies the copolarized signal component with the pump but keeps the orthogonally polarized one almost unchanged. Because of this imbalance, SRS rotates **S** toward **P** as dictated by Eq. (2.32). However, PMD scatters the signal SOP away from the pump. If PMD is relatively large, θ changes so fast that the signal experiences only an average local gain everywhere, and the accumulative fluctuations of the output signal are small. Also, if PMD is negligible, θ changes almost deterministically, and the signal fluctuations are again small. However, when the effective fiber length is comparable with the PMD diffusion length, the signal experiences random gain from section to section, resulting in large signal fluctuations.

To illustrate the effect of PMD on the performance of Raman amplifiers, we focus on a 10-km-long amplifier pumped with 1 W of power using a 1.45-μm laser. The 1.55-μm signal is assumed to be located at the Raman gain peak ($\Omega_R/2\pi = 13.2$ THz). The Raman gain coefficients have values $g_1 = 0.60\,W^{-1}/km$ and $g_2 = 0.0071\,W^{-1}/km$.[5,9] Fiber losses are taken to be 0.273 and 0.2 dB/km at the pump and the signal wavelengths, respectively. Figure 1 shows how the average gain and σ_s change with the PMD parameter D_p when the input signal is copolarized (solid curves) or orthogonally polarized (dashed curves) to the pump. The curves are shown for both the forward- and the backward-pumping schemes. When D_p is zero, the two beams maintain their SOPs, and the copolarized signal experiences a maximum gain of 17.6 dB but the orthogonally polarized one has a 1.7-dB loss, regardless of the pumping configuration. The loss is not exactly 2 dB because a small gain exists for the orthogonally polarized input signal $(4ag_2P_{in}L_{eff}/3)$. As PMD increases, the gain difference between the copolarized and the orthogonally polarized cases decreases and disappears eventually.

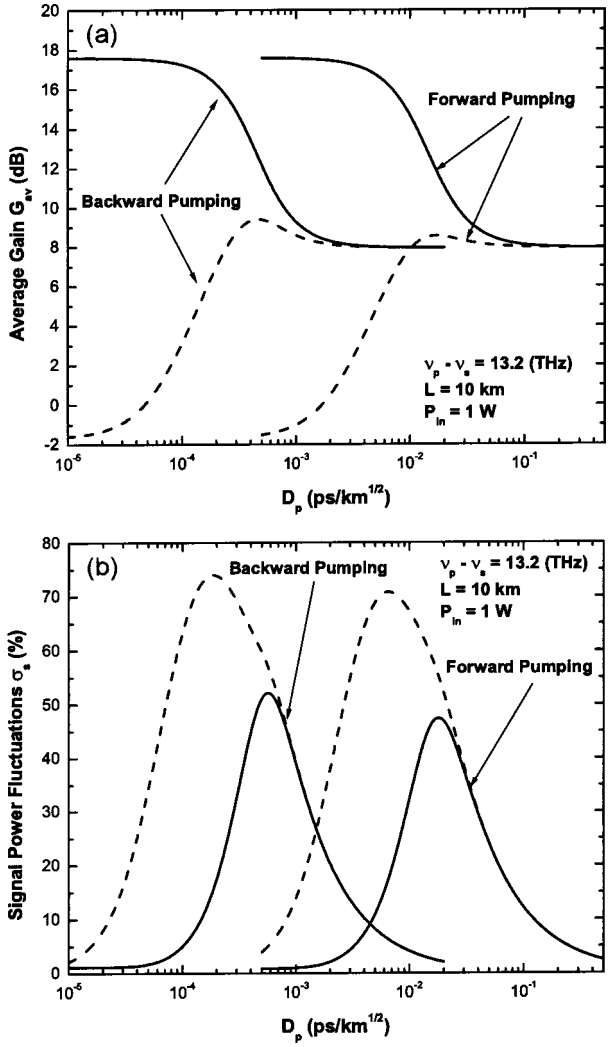

Fig. 1. (a) Average gain and (b) standard deviation of signal fluctuations at the output of a Raman amplifier as a function of the PMD parameter for forward and backward pumping. The solid and dashed curves correspond to copolarized and orthogonally polarized signals, respectively.

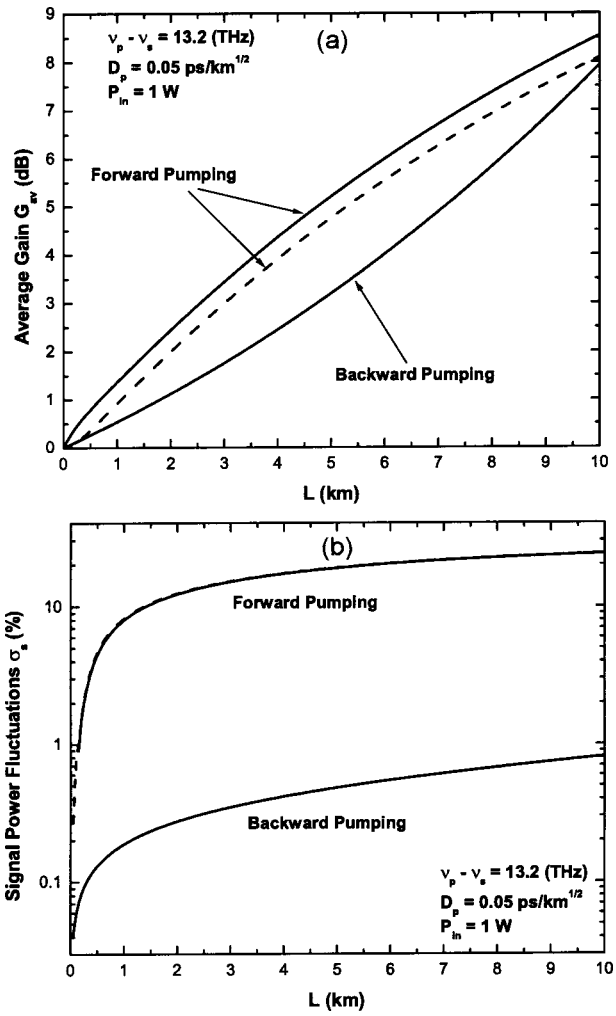

Fig. 2. (a) Average gain and (b) level of signal fluctuations as a function of amplifier length for a fiber with $D_p = 0.05\ \mathrm{ps}/\sqrt{\mathrm{km}}$. The solid and dashed curves correspond to copolarized and orthogonally polarized signals, respectively. The two curves nearly coincide in the case of backward pumping.

The level of signal fluctuations in Fig. 1 increases quickly with the PMD parameter, reaches a peak, and then decreases slowly to zero with further increase in D_p. The location of the peak depends on the pumping scheme as well as on the initial polarization of the pump. The noise level can exceed 20% for $D_p = 0.05\ \mathrm{ps}/\sqrt{\mathrm{km}}$ in the case of forward pumping. If a fiber with low PMD is used, the noise level can exceed 70% under some conditions. These results suggest that forward-pumped Raman amplifiers will perform better if a fiber with $D_p > 0.1\ \mathrm{ps}/\sqrt{\mathrm{km}}$ is used. The curves for backward pumping are similar to those for forward pumping but shift to smaller D_p values and have a higher peak. In spite of an enhanced peak, the backward pumping produces the least fluctuations for all fibers for which $D_p > 0.01\ \mathrm{ps}/\sqrt{\mathrm{km}}$.

Note in Fig. 1 that the curves in the case of backward pumping are nearly identical to those for forward pumping except that they are shifted to the left. As a result, the solid and dashed curves merge at a value of D_p that is smaller by a factor of approximately 30. This difference is related to the definition of $\Omega_R = \xi\omega_p - \omega_s$ in Eq. (2.32). In the case of backward pumping, $|\Omega_R| = \omega_p + \omega_s$ is approximately 30 times larger than the value of $\Omega_R = \omega_p - \omega_s$ in the forward-pumping case.

In practice, fibers used to make a Raman amplifier have a constant value of D_p. Figure 2 shows the average Raman gain and σ_s as a function of amplifier length for a fiber with $D_p = 0.05\ \mathrm{ps}/\sqrt{\mathrm{km}}$. All the other parameters are the same as in Fig. 1. The solid and dashed curves correspond to copolarized and orthogonally polarized cases, respectively (the two curves are indistinguishable in the case of backward pumping). Physically, it takes some distance for the orthogonally polarized signal to adjust its SOP through PMD before it can experience the full Raman gain. Within the PMD diffusion length (around 175 m in this case of forward pumping), fiber loss dominates and the signal power decreases; beyond the diffusion length, Raman gain dominates and the signal power increases. The gain difference seen in Fig. 1 between the copolarized and orthogonally polarized cases comes from this initial difference. In the case of backward pumping, the PMD diffusion length becomes so

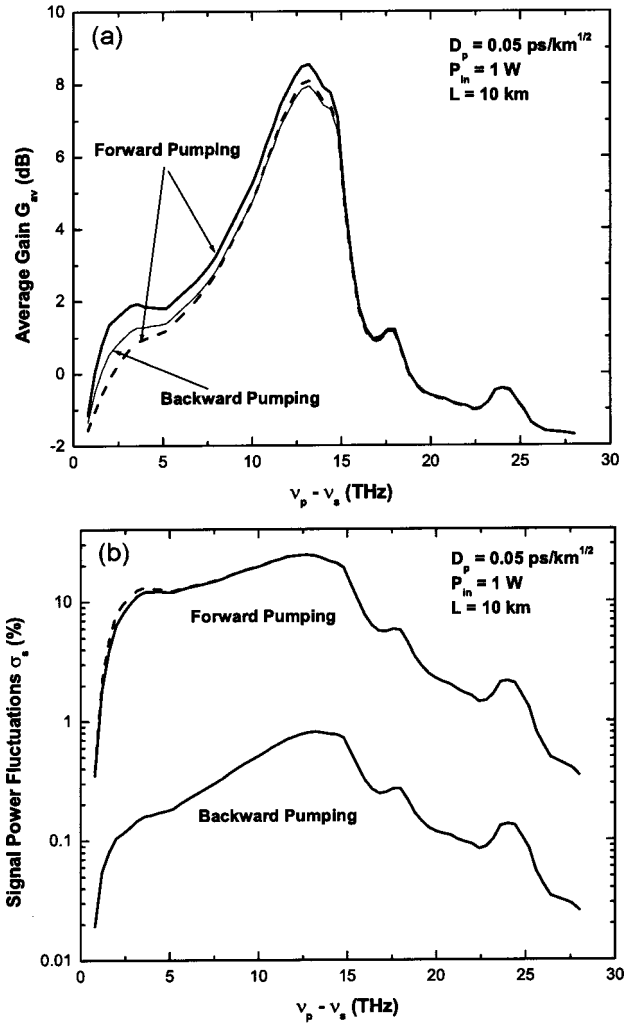

Fig. 3. (a) Average gain and (b) level of signal fluctuations plotted as a function of pump-signal detuning. The solid and dashed curves correspond to copolarized and orthogonally polarized cases, respectively. The thin solid curve is for a backward-pumped Raman amplifier and does not change much with signal polarization.

small (approximately 0.2 m) that this difference completely disappears. The level of signal fluctuations depends strongly on the relative directions of pump and signal propagation. In the case of forward pumping, σ_s grows monotonically with the distance, reaching 24% at the end of the 10-km fiber. In contrast, σ_s is only 0.8% even for a 10-km-long amplifier in the case of backward pumping, a value 30 times smaller than that occurring in the forward-pumping case. The curves are almost identical for all input signal SOPs when $D_p = 0.05\ \mathrm{ps}/\sqrt{\mathrm{km}}$.

So far we assumed that the signal wavelength coincided with the Raman-gain peak. Figure 3 shows the effect of pump-signal detuning for a 10-km-long amplifier under the same conditions. The frequency dependence of Raman gain coefficients was taken from Refs. 5 and 9. In the case of forward pumping, the average Raman gain is different for copolarized (solid curve) and orthogonally polarized (dashed curve) signals, and the difference is larger when the signal frequency is close to the pump. However, the difference disappears when the signal frequency deviates more than 15 THz from the pump because of increased PMD effects. In the case of backward pumping, PMD effects are so huge over the whole spectrum that the differences disappear completely, and the same spectrum (thin curve) is obtained for all input SOPs of the signal. Signal fluctuations depend on the PMD parameter as well as on the Raman gain. The larger the gain, the larger the fluctuations. For this reason, σ_s is maximum at the Raman-gain peak. Again, fluctuations in the case of backward pumping are 30 times smaller than those in the case of forward pumping because of the ratio $(\omega_p + \omega_s)/(\omega_p - \omega_s) \approx 30$.

4. PROBABILITY DISTRIBUTION OF THE AMPLIFIED SIGNAL

The moment method used in Section 3 to obtain the average and variance of the amplified signal becomes increasingly complex for higher-order moments. It will be much better if we can determine the probability distribution of the amplified signal because it contains, by definition, all the statistical information. To this end, we begin by finding the instantaneous gain of the Raman amplifier.

The fluctuating amplifier gain $G(L)$ for a Raman amplifier of length L can be found from Eqs. (2.30)–(2.32) as (in decibels)

$$G_{\mathrm{dB}} = a \ln\left[\frac{S_0(L)}{S_0(0)}\right] = a\left[\frac{g_1}{2}(1 + 3\mu)P_{\mathrm{in}}L_{\mathrm{eff}} - \alpha_s L\right] + \frac{a}{2}g_R\int_0^L P_0(z)[\hat{p}(z)\cdot\hat{s}(z)]\mathrm{d}z, \qquad (4.1)$$

where $a = 10/\ln(10) = 4.343$ and $\hat{s}$ is the unit vector in the direction of $\mathbf{s}$. Using $\mathbf{s} = s_0\hat{s}$ in Eq. (2.32), $\hat{s}$ is found to satisfy

$$\frac{\mathrm{d}\hat{s}}{\mathrm{d}z} = \frac{g_R}{2}P_0(z)[\hat{p} - (\hat{p}\cdot\hat{s})\hat{s}] - \Omega_R\mathbf{b}\times\hat{s}. \qquad (4.2)$$

In this equation $\hat{p}$ represents the pump polarization at the input end. Thus, $\hat{s}$ becomes random only because of birefringence fluctuations. If polarization scrambling is used to randomize $\hat{p}$, $\hat{p}$ and $\hat{s}$ both become random (see Section 6). However, it is only the scalar product $\hat{p}\cdot\hat{s} \equiv \cos\theta$ that determines G_{dB}.

To find the probability distribution of G_{dB}, we note that the second term in Eq. (4.1) can be written as $\Sigma_{i=1}^{N}P_0(z_i)\cos[\theta(z_i)]\Delta z$ if we divide fiber length L into N segments of length Δz. Thus the random variable G_{dB} is formed from a sum of many random variables with identical statistics. According to the central limit theorem,[24] the probability density of $G(L)$ should be Gaussian as long as the correlation between $\cos[\theta(z)]$ and $\cos[\theta(z')]$ goes to zero sufficiently rapidly as $|z - z'|$ increases, no matter what the statistics of $\hat{p}(z)$ and $\hat{s}(z)$ are. In practice, we expect this correlation to decay exponentially over a length of PMD diffusion length L_d. For fiber lengths $L \gg L_d$, we thus expect G_{dB} to have a Gaussian distribution. Such a Gaussian distribution of G_{dB} has been observed experimentally in Ref. 25. Our vector theory explains this experimental result in a simple way.

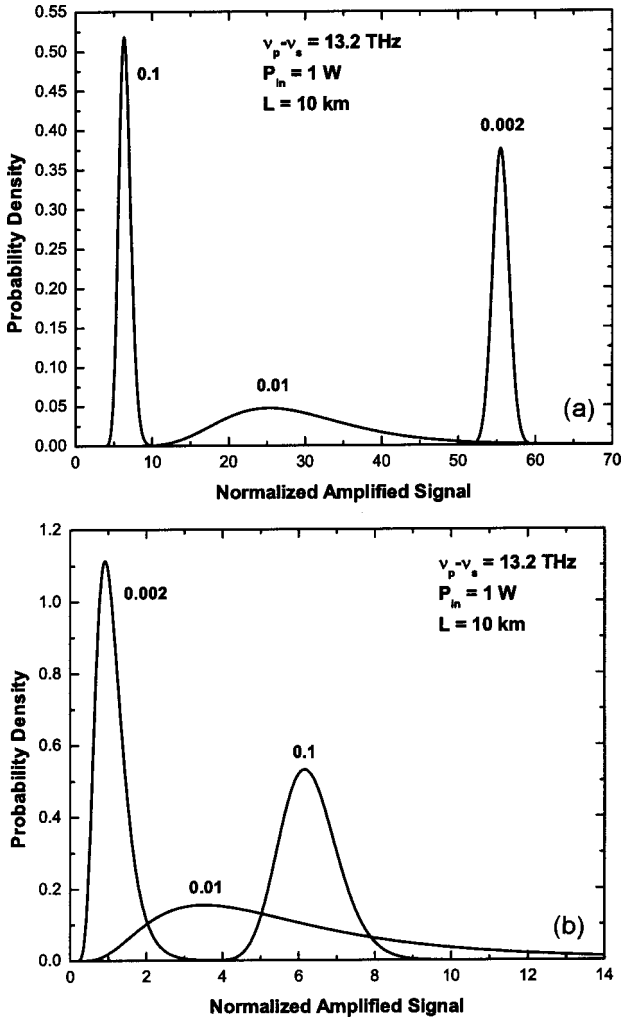

Fig. 4. Probability density of an amplified signal for three values of D_p (D_p in units of $\mathrm{ps}/\sqrt{\mathrm{km}}$) in the cases of (a) copolarized and (b) orthogonally polarized signals. The amplified signal is normalized to the input signal power.

It is clear from Eq. (4.1) that $\ln[S_0(L)]$ will also follow a Gaussian distribution. As a result, the probability distribution of the amplified signal power, $S_0(L)$, at the amplifier output corresponds to a log-normal distribution[26] and can be written as

$$p[S_0(L)] = \frac{[\ln(\sigma_s^2 + 1)]^{-1/2}}{S_0(L)\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\ln(\sigma_s^2 + 1)} \times \ln^2\left[\frac{S_0(L)\sqrt{\sigma_s^2 + 1}}{\langle S_0(L)\rangle} \right] \right\}, \tag{4.3}$$

where σ_s^2 is defined in Eq. (3.1) and can be calculated explicitly as discussed in Section 3.

Figure 4 shows how the probability density changes with the PMD parameter in the cases of copolarized and orthogonally polarized input signals, respectively, under the conditions of Fig. 1. When the PMD effects are relatively small, the two distributions are relatively narrow and are centered at quite different locations in the copolarized and orthogonally polarized cases. When $D_p = 0.01\,\text{ps}/\sqrt{\text{km}}$, the two distributions broaden considerably and begin to approach each other. For larger values of the PMD parameter $D_p = 0.1\,\text{ps}/\sqrt{\text{km}}$, they become narrow again and their peaks almost overlap because the amplified signal becomes independent of the input SOP.

5. POLARIZATION-DEPENDENT GAIN

Similar to the importance of the concept of differential group delay in describing the PMD effects on pulse propagation, the PMD effects in Raman amplifiers can be quantified by use of the concept of PDG, a quantity defined as the difference between the maximum and the minimum values of G while varying the SOP of the input signal. The gain difference $\Delta = G_{\max} - G_{\min}$ is itself random because both $G_{\max}$ and $G_{\min}$ are random. It is important to know the statistics of Δ and its relationship to the operating parameters of a Raman amplifier, because they can identify the conditions under which the PDG can be reduced to acceptable low levels. In this section we introduce a PDG vector and use it to describe the statistics of PDG.

The polarization-dependent loss is often described by introducing a polarization-dependent loss vector.[27,28] The same technique can be used for the PDG in Raman amplifiers. The PDG vector $\mathbf{\Delta}$ is introduced in such a way that its magnitude gives the PDG value Δ (in decibels), but its direction coincides with the direction of $\mathbf{S}(L)$ for which the gain is maximum. From Eqs. (2.31) and (2.32), the dynamic equation for $\mathbf{\Delta}$ is found to be

$$\frac{d\mathbf{\Delta}}{dz} = \frac{g_R}{2}\Delta \coth\left(\frac{\Delta}{2a}\right)[\mathbf{P} - (\mathbf{P}\cdot\hat{\Delta})\hat{\Delta}] + ag_R(\mathbf{P}\cdot\hat{\Delta})\hat{\Delta} - \Omega_R\mathbf{b}\times\mathbf{\Delta}. \tag{5.1}$$

where $\hat{\Delta}$ is the unit vector in the direction of $\mathbf{\Delta}$. Appendix C provides details of the derivation of this equation. If the PDG is not too large, $\Delta\coth(\Delta/2a)$ can be expanded in a Taylor series as

$$\Delta\coth\left(\frac{\Delta}{2a}\right) \approx 2a + \frac{\Delta^2}{6a}. \tag{5.2}$$

Keeping only the linear terms in Δ, Eq. (5.1) reduces to the following linear Langevin equation:

$$\frac{d\mathbf{\Delta}}{dz} = ag_R\mathbf{P} - \Omega_R\mathbf{b}\times\mathbf{\Delta}. \tag{5.3}$$

The validity of this linearized equation depends on the conditions under which the Δ^2 term in Eq. (5.2) can be neglected. The validity condition can be written as $\Delta_{\max} \ll \sqrt{12}a \approx 15\,\text{dB}$. This requirement is satisfied in practice for most Raman amplifiers.

Equation (5.3) can be readily solved because of its linear nature. The solution is given by

$$\mathbf{\Delta}(L) = ag_R\vec{R}(L)\int_0^L \vec{R}^{-1}(z)\mathbf{P}(z)dz, \tag{5.4}$$

where the PMD-induced rotation matrix $\vec{R}(z)$ is obtained from $d\vec{R}/dz = -\Omega_R\mathbf{b}\times\vec{R}$. For fibers much longer than the birefringence correlation length, the dynamics of $\mathbf{\Delta}$ corresponds to that of the Brownian motion in three dimensions [see Ref. 29, where Eq. (5.4) is used for the PMD vector]. As a result, $\mathbf{\Delta}$ follows a three-dimensional Gaussian distribution.

The moments of $\mathbf{\Delta}$ can be obtained from Eq. (5.3) using the same procedure as that used in the last section for calculating the average gain and signal fluctuations.[24] In Appendix D we provide the details. As shown there, the first two moments and the covariance matrix $\vec{C}$ defined as $\vec{C} = \langle\mathbf{\Delta}\mathbf{\Delta}\rangle - \langle\mathbf{\Delta}\rangle\langle\mathbf{\Delta}\rangle$ satisfy

$$\frac{d\langle\Delta^2\rangle}{dz} = 2ag_RP_0(z)\hat{p}\cdot\langle\mathbf{\Delta}\rangle, \tag{5.5}$$

$$\frac{d\langle\mathbf{\Delta}\rangle}{dz} = -\eta\langle\mathbf{\Delta}\rangle + ag_RP_0(z)\hat{p}, \tag{5.6}$$

$$\frac{d\vec{C}}{dz} = -3\eta\vec{C} + \eta[\langle\Delta^2\rangle\vec{I} - \langle\mathbf{\Delta}\rangle\langle\mathbf{\Delta}\rangle]. \tag{5.7}$$

Equations (5.5) and (5.6) can be easily integrated over fiber length L. They provide the following analytical results in the case of forward pumping:

$$\langle\mathbf{\Delta}\rangle = \frac{ag_RP_{\text{in}}\hat{p}}{\eta - \alpha_p}[1 - \alpha_pL_{\text{eff}} - \exp(-\eta L)], \tag{5.8}$$

$$\langle\Delta^2\rangle = \frac{2(ag_RP_{\text{in}})^2}{\eta^2 - \alpha_p^2}[(1 - \alpha_pL_{\text{eff}})\exp(-\eta L) - 1 + (\alpha_p + \eta)L_{\text{eff}}(1 - \alpha_pL_{\text{eff}}/2)]. \tag{5.9}$$

In the case of backward pumping, $\langle\mathbf{\Delta}\rangle$ and $\langle\Delta^2\rangle$ are also given by these equations provided α_p is replaced with $-\alpha_p$, P_{in} is replaced with $P_{\text{in}}\exp(-\alpha_pL)$, L_{eff} is redefined as $[\exp(\alpha_pL) - 1]/\alpha_p$, and $\Omega_R = \omega_p - \omega_s$ in the expression of η is replaced by $\Omega_R = -(\omega_p + \omega_s)$. An analytical expression for $\vec{C}$ can also be obtained by integrating Eq. (5.7).

We now consider the probability distribution of the PDG vector. It is convenient to choose $\hat{p}$ along an axis of the Stoke space, say $\hat{p} = \hat{e}_1$, because matrix $\vec{C}$ is then diagonal. The probability density function of $\mathbf{\Delta} \equiv \Delta_1\hat{e}_1 + \Delta_2\hat{e}_2 + \Delta_3\hat{e}_3$ in this case can be written as

$$p(\mathbf{\Delta}) = \frac{(2\pi)^{-3/2}}{\sigma_\parallel \sigma_\perp^2} \exp\left[-\frac{(\Delta_1 - \Delta_0)^2}{2\sigma_\parallel^2} - \frac{\Delta_2^2 + \Delta_3^2}{2\sigma_\perp^2}\right], \tag{5.10}$$

where $\Delta_0 = |\langle \mathbf{\Delta} \rangle|$ whereas $\sigma_\parallel^2$ and $\sigma_\perp^2$ are the variances of the PDG vector in the parallel and perpendicular directions of $\hat{p}$, respectively. Both can be found from Eq. (5.7) in an analytical form as

$$\sigma_\parallel^2 = \eta \int_0^L [\langle \Delta^2 \rangle - \langle \mathbf{\Delta} \rangle^2] \exp[-3\eta(L - z)]\mathrm{d}z, \tag{5.11}$$

$$\sigma_\perp^2 = \eta \int_0^L \langle \Delta^2 \rangle \exp[-3\eta(L - z)]\mathrm{d}z. \tag{5.12}$$

These equations show that $\sigma_\parallel < \sigma_\perp$ when PMD is small because the pump mostly amplifies the copolarized signal. When the effective fiber length L_{eff} is much larger than the PMD diffusion length L_d, $\langle \Delta^2 \rangle \gg \langle \mathbf{\Delta} \rangle^2 \approx 0$ and $\sigma_\parallel \approx \sigma_\perp$. In this case, $p(\mathbf{\Delta})$ becomes a uniform three-dimensional Gaussian distribution centered at zero.

In practice, one is often interested in the statistics of the PDG magnitude Δ. Its probability density can be found from Eq. (5.10) after writing $\mathbf{\Delta}$ in spherical coordinates and integrating over the two angles. The result is found to be

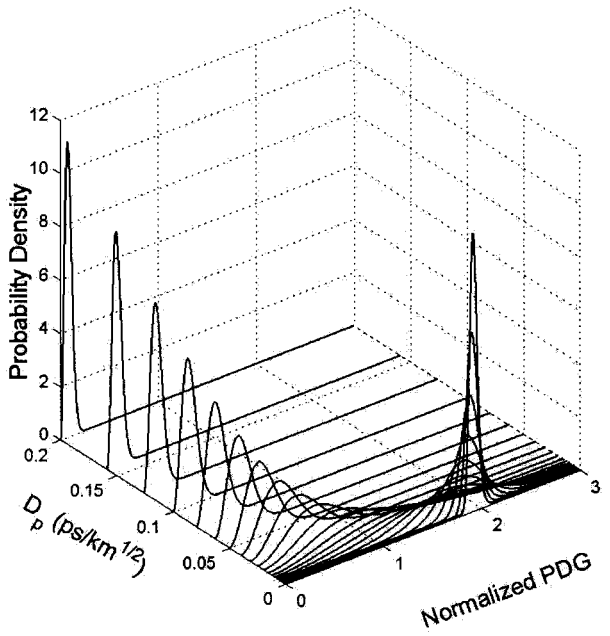

Fig. 5. Probability distribution of PDG as a function of D_p under conditions of Fig. 1. The PDG value is normalized to the average gain G_{av}.

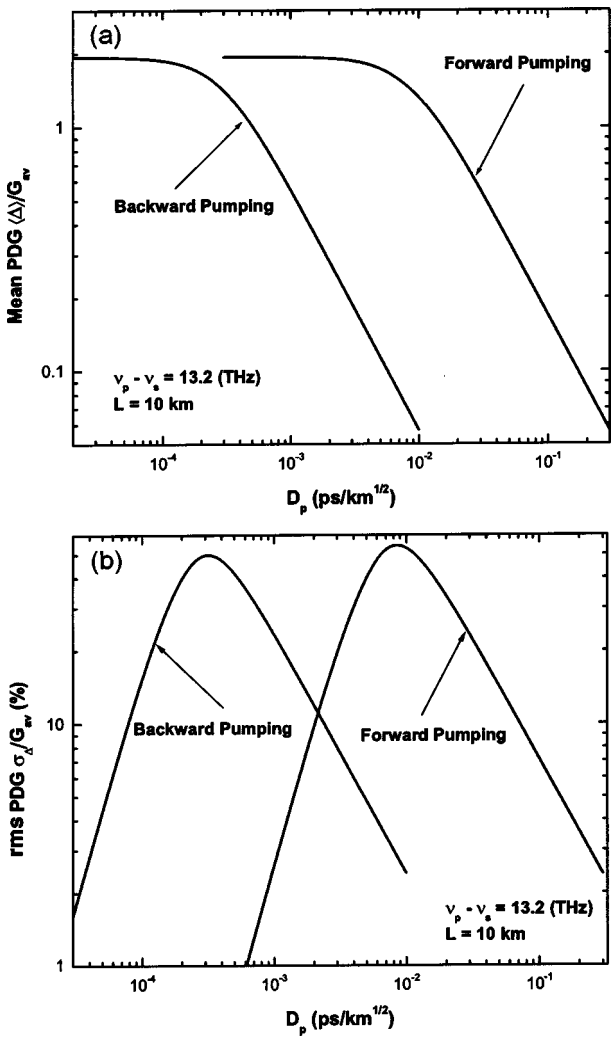

Fig. 6. (a) Mean PDG and (b) variance σ_Δ (both normalized to the average gain G_{av}) as a function of PMD parameter under forward- and backward-pumping conditions.

$$p(\Delta) = \frac{\Delta}{2\sigma_\parallel \sigma} \exp\left[-\frac{\Delta^2(r-1) - r\Delta_0^2}{2\sigma^2}\right] \times \left\{ \mathrm{erf}\left[\frac{\Delta(r-1) + r\Delta_0}{\sqrt{2}\sigma}\right] + \mathrm{erf}\left[\frac{\Delta(r-1) - r\Delta_0}{\sqrt{2}\sigma}\right] \right\}, \tag{5.13}$$

where $\sigma^2 = \sigma_\perp^2(r-1)$, $r = \sigma_\perp^2/\sigma_\parallel^2$, and the error function is defined as $\mathrm{erf}(x) = (2/\sqrt{\pi})\int_0^x \exp(-y^2)\mathrm{d}y$. Figure 5 shows how $p(\Delta)$ changes with D_p in the case of forward pumping. All the parameters are the same as in Fig. 1. The PDG values are normalized to the average gain $G_{\text{av}} = ag_1(1 + 3\mu)P_{\text{in}}L_{\text{eff}}/2$ (in decibels) so that the curves are pump-power independent. In the limit $D_p \to 0$, $p(\Delta)$ becomes a delta function located at the maximum gain difference (almost $2G_{\text{av}}$) as little gain exists for the orthogonally polarized signal. As D_p increases, $p(\Delta)$ broadens quickly because PMD changes the signal SOP randomly. If D_p is relatively small, diffusion

length L_d is larger than or comparable with the effective fiber length L_{eff}, and $p(\Delta)$ remains centered at almost the same location but broadens because of large fluctuations. Its shape mimics a Gaussian distribution. When D_p is large enough that $L_{\text{eff}} \gg L_d$, $p(\Delta)$ becomes Maxwellian and its peak shifts to smaller values. This is the behavior observed experimentally in Ref. 12.

The mean value of PDG, $\langle\Delta\rangle$, and the variance of PDG fluctuations, $\sigma_\Delta^2 = \langle\Delta^2\rangle - \langle\Delta\rangle^2$, can be found by use of the PDG distribution in Eq. (5.13). Figure 6 shows how these two quantities vary with the PMD parameter for the same Raman amplifier used for Fig. 1. Both $\langle\Delta\rangle$ and σ_Δ are normalized to the average gain G_{av}. As expected, the mean PDG decreases monotonically as D_p increases. The mean PDG $\langle\Delta\rangle$ is not exactly $2G_{\text{av}}$ when $D_p = 0$ because the gain is not zero when pump and signal are orthogonally polarized. Note however that $\langle\Delta\rangle$ can be as large as 30% of the average gain for $D_p = 0.05\,\text{ps}/\sqrt{\text{km}}$ in the case of forward pumping and it decreases slowly with D_p after that, reaching a value of 8% for $D_p = 0.2\,\text{ps}/\sqrt{\text{km}}$. This is precisely what was observed in Ref. 12 through experiments and numerical simulations.

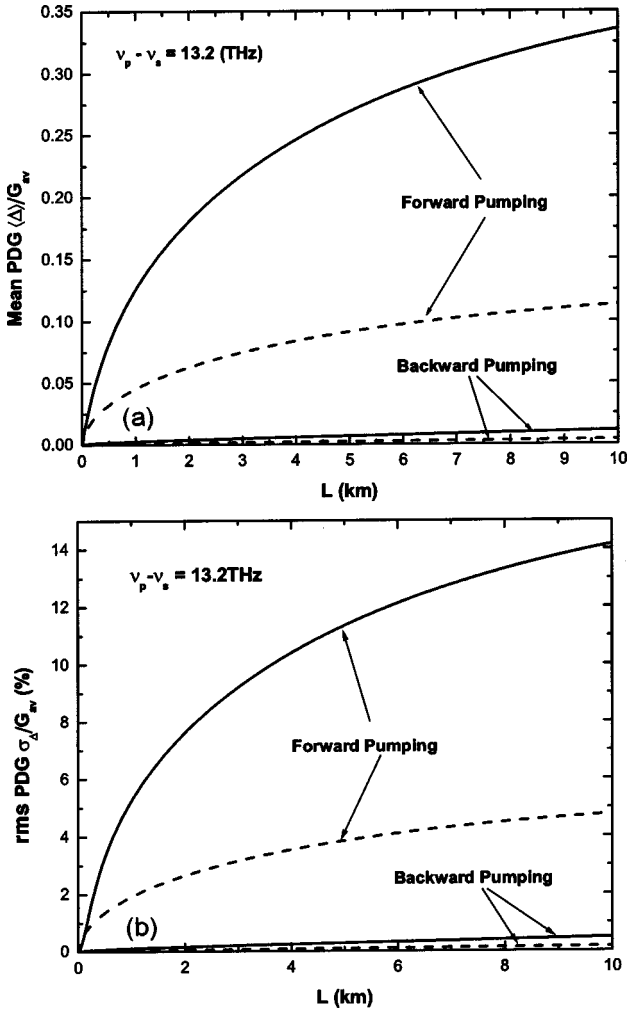

Fig. 7. (a) Mean PDG and (b) variance σ_Δ (both normalized to the average gain G_{av}) as a function of amplifier length under forward and backward pumping. The solid and dashed curves correspond to $D_p = 0.05$ and $0.15\,\text{ps}/\sqrt{\text{km}}$, respectively.

In the case of backward pumping, the behavior is nearly identical to that in the case of forward pumping except that the curve shifts to a value of D_p smaller by a factor of approximately 30.

As seen in Fig. 6(b), the rms value of PDG fluctuations increases rapidly as D_p becomes nonzero, peaks to a value close to 56% of G_{av} for D_p near $0.01\,\text{ps}/\sqrt{\text{km}}$ in the case of forward pumping, and then begins to decrease. Again, fluctuations can exceed 7% of the average gain level even for $D_p = 0.1\,\text{ps}/\sqrt{\text{km}}$. A similar behavior holds for backward pumping, as seen in Fig. 6. Both the mean and the rms values of PDG fluctuations decrease with D_p inversely for $D_p > 0.03\,\text{ps}/\sqrt{\text{km}}$ ($L_d < 0.5\,\text{km}$ for $\Omega_R/2\pi = 13.2\,\text{THz}$). This D_p^{-1} dependence is in good agreement with the experimental results in Ref. 12.

The D_p^{-1} dependence of $\langle\Delta\rangle$ and σ_Δ can be deduced analytically from Eq. (5.13) in the limit $L_{\text{eff}} \gg L_d$. In this limit, the PDG distribution $p(\Delta)$ becomes approximately Maxwellian and the average and rms values of PDG in the case of forward pumping are given by

$$\langle\Delta\rangle \approx \frac{4ag_R P_{\text{in}}}{\sqrt{\pi} D_p |\Omega_R|}[L_{\text{eff}}(1 - \alpha_p L_{\text{eff}}/2)]^{1/2}, \quad (5.14)$$

$$\sigma_\Delta \approx [(3\pi/8 - 1)]^{1/2}\langle\Delta\rangle. \quad (5.15)$$

The same equations hold in the case of backward pumping except that $|\Omega_R| = \omega_p - \omega_s$ should be replaced with $\omega_p + \omega_s$. Figure 7 shows the mean PDG and the rms value of PDG as a function of propagation distance for a Raman amplifier with $D_p = 0.05\,\text{ps}/\sqrt{\text{km}}$. All the parameters are the same as in Fig. 2. Both $\langle\Delta\rangle$ and σ_Δ are negligibly small in the case of backward pumping, indicating the advantages of such a pumping scheme.

6. POLARIZATION SCRAMBLING OF THE PUMP

The technique of polarization scrambling is sometimes used to reduce the PDG in Raman amplifiers.[30,31] In this technique, the pump SOP is changed randomly as the signal is amplified so that the signal undergoes different local gain in different parts of the fiber, resulting effectively in an average gain that is independent of the signal SOP. The theory developed in Section 5 can be used to find the dependence of the PDG on the degree of polarization (DOP) of the pump wave.

Polarization scrambling does not change the total power of pump as only its SOP is changed randomly. If we assume that $\hat{p}(z)$ is a stationary stochastic process, its correlation at two points within the fiber can be written as

$$\langle[\hat{p}(z_1) - \langle\hat{p}\rangle][\hat{p}(z_2) - \langle\hat{p}\rangle]\rangle \equiv \overleftrightarrow{\Gamma}(z_2 - z_1) = \overleftrightarrow{\Gamma}_0\Gamma(z_2 - z_1), \quad (6.1)$$

where the correlation matrix $\overleftrightarrow{\Gamma}(z_2 - z_1)$ is assumed to be separable as shown in Eq. (6.1). The DOP of the pump is related to the average of $\hat{p}$ as $d_p = |\langle\hat{p}\rangle|$.[32] The pump DOP d_p, the covariance matrix $\overleftrightarrow{\Gamma}_0$, and the correlation function $\Gamma(z_2 - z_1)$ depend on how scrambling rotates $\hat{p}$ randomly on the Poincaré sphere. When the pump is

completely polarized (no scrambling), $d_p = 1$ and $\vec{\vec{\Gamma}}_0 = 0$; in the opposite limit in which the pump is completely unpolarized, $d_p = 0$ and $\vec{\vec{\Gamma}}_0 = \vec{\vec{I}}/3$. The form of $\Gamma(z_2 - z_1)$ depends on the specific scrambling technique used in practice. In the following discussion, we assume it to vary as $\Gamma(z_2 - z_1) = \exp(-\gamma_c|z_2 - z_1|)$, where $\gamma_c = 1/l_c$ and l_c is the correlation length. In practice, l_c depends on the coherence length of the pump source and is typically ~1 m.

Because of the randomness of the pump polarization, the XPM-induced NPR term in Eq. (2.26) becomes random as well. However, correlation length l_c is so small compared with the NPR beat length (~10 km) that the NPR term contributes only from the rotation around the average pump polarization $\langle\hat{p}\rangle$. The NPR induced by the pump polarization fluctuations is negligible if we replace the deterministic transformation in Eq. (2.27) with

$$\mathbf{V} = \exp\left\{-\varepsilon_{sp}\left[\int_0^z P_0(z)\mathrm{d}z\right]\langle\hat{p}\rangle\times\right\}\mathbf{V}', \tag{6.2}$$

where $\mathbf{V}$ is an arbitrary vector in the Stokes space (Appendix E provides more details). Therefore, Eqs. (2.31) and (2.32) remain valid even in the presence of pump polarization scrambling.

As before, the PDG vector satisfies Eq. (5.1). This equation can again be simplified to obtain Eq. (5.3) with the only difference that now both $\hat{p}$ and $\mathbf{b}$ are random. The solution is still given by Eq. (5.4). More importantly, the probability distribution of $\mathbf{\Delta}$ remains Gaussian (in all three dimensions) as long as fiber length L is much longer than the correlation length associated with pump polarization scrambling. However, as the pump polarization varies with z, the solution of Eqs. (5.5) and (5.6) is given by

$$\langle\mathbf{\Delta}(L)\rangle_b = ag_R\int_0^L P_0(z)\hat{p}(z)\exp[-\eta(L-z)]\mathrm{d}z, \tag{6.3}$$

$$\begin{aligned}\langle\Delta^2(L)\rangle_b &= 2ag_R\int_0^L P_0(z)\hat{p}(z)\cdot\langle\mathbf{\Delta}(z)\rangle_b\mathrm{d}z,\\ &= 2(ag_R)^2\int_0^L\mathrm{d}z_1\int_0^{z_1}\mathrm{d}z_2P_0(z_1)P_0(z_2)\hat{p}(z_1)\\ &\quad\cdot\hat{p}(z_2)\exp[-\eta(z_1-z_2)],\end{aligned} \tag{6.4}$$

where the subscript b denotes the average over the birefringence fluctuations. Since $\hat{p}(z)$ is random as well, we need to average Eqs. (6.3) and (6.4) over the pump SOP. In the case of forward pumping, the final analytic expressions are found to be

$$\langle\langle\mathbf{\Delta}(L)\rangle_b\rangle_p = \frac{ag_RP_{\mathrm{in}}\langle\hat{p}\rangle}{\eta-\alpha_p}[1-\alpha_pL_{\mathrm{eff}}-\exp(-\eta L)], \tag{6.5}$$

$$\begin{aligned}\langle\langle\Delta^2(L)\rangle_b\rangle_p = &\frac{2(ag_RP_{\mathrm{in}})^2\mathrm{Tr}(\vec{\vec{\Gamma}}_0)}{(\gamma_c+\eta)^2-\alpha_p^2}\\ &\times\{(1-\alpha_pL_{\mathrm{eff}})\exp[-(\gamma_c+\eta)L]-1\\ &+L_{\mathrm{eff}}(\gamma_c+\eta+\alpha_p)(1-\alpha_pL_{\mathrm{eff}}/2)\}\\ &+\frac{2(ag_RP_{\mathrm{in}})^2d_p^2}{\eta^2-\alpha_p^2}\{(1-\alpha_pL_{\mathrm{eff}})\exp(-\eta L)\\ &-1+L_{\mathrm{eff}}(\eta+\alpha_p)(1-\alpha_pL_{\mathrm{eff}}/2)\},\end{aligned} \tag{6.6}$$

where the subscript p denotes average over the ensemble of pump polarization and Tr stands for the trace. From Eq. (6.1), $\mathrm{Tr}(\vec{\vec{\Gamma}}_0)$ is related to the DOP of pump as $\mathrm{Tr}(\vec{\vec{\Gamma}}_0) = 1 - d_p^2$.

The covariance matrix $\vec{\vec{C}}$ as defined in Section 5 can also be found by use of the same technique and requires averaging over both ensembles of birefringence and pump polarization. It satisfies the following equation (see Appendix D for details):

$$\frac{\mathrm{d}\vec{\vec{C}}}{\mathrm{d}z} = -3\eta\vec{\vec{C}} + \eta[\langle\!\langle\Delta^2\rangle\!\rangle\vec{\vec{I}} - \langle\!\langle\mathbf{\Delta}\rangle\!\rangle\langle\!\langle\mathbf{\Delta}\rangle\!\rangle] + H_p, \tag{6.7}$$

where the last term depends on polarization scrambling and is found to be

$$\begin{aligned}H_p &= ag_RP_0(z)[\langle\hat{p}(z)\langle\mathbf{\Delta}(z)\rangle_b\rangle_p + \langle\langle\mathbf{\Delta}(z)\rangle_b\hat{p}(z)\rangle_p\\ &\quad-\langle\hat{p}\rangle\langle\!\langle\mathbf{\Delta}\rangle\!\rangle - \langle\!\langle\mathbf{\Delta}\rangle\!\rangle\langle\hat{p}\rangle].\\ &= \frac{2(ag_RP_{\mathrm{in}})^2\vec{\vec{\Gamma}}_0}{\gamma_c+\eta-\alpha_p}\{\exp(-2\alpha_pz)\\ &\quad-\exp[-(\gamma_c+\eta+\alpha_p)z]\}.\end{aligned} \tag{6.8}$$

If the pump polarization is not scrambled, $d_p = 1$ and $\vec{\vec{\Gamma}}_0 = 0$. In that case Eqs. (6.5)–(6.7) reduce to Eqs. (5.7)–(5.9), as expected. The same results hold in the case of backward pumping except that α_p is replaced with $-\alpha_p$, P_{in} is replaced with $P_{\mathrm{in}}\exp(-\alpha_pL)$, L_{eff} is redefined as $[\exp(\alpha_pL)-1]/\alpha_p$, and $\Omega_R = \omega_p - \omega_s$ in the expression of η is replaced by $\Omega_R = -(\omega_p + \omega_s)$.

It is evident from Eqs. (6.5)–(6.7) that scrambling affects the PDG considerably. Although the results are quite complicated when correlation length l_c and effective fiber length L_{eff} are comparable, they can be simplified considerably in a practical situation in which $L_{\mathrm{eff}} \gg l_c$. This is almost always the case since l_c is typically ~1 m whereas L_{eff} is likely to exceed 1 km. In this limit, H_p in Eq. (6.8) reduces to

$$H_p \approx 2(ag_RP_{\mathrm{in}})^2\vec{\vec{\Gamma}}_0l_c(1-\alpha_pL_{\mathrm{eff}})^2, \tag{6.9}$$

and it becomes negligible in Eq. (6.7) for $l_c \ll L_{\mathrm{eff}}$. In the same limit, the first term in Eq. (6.6) becomes much smaller than the second term and $\langle\!\langle\Delta^2(L)\rangle\!\rangle$ is given by

$$\begin{aligned}\langle\langle\Delta^2(L)\rangle_b\rangle_p \approx &\frac{2(ag_RP_{\mathrm{in}}d_p)^2}{\eta^2-\alpha_p^2}[(1-\alpha_pL_{\mathrm{eff}})\exp(-\eta L)-1\\ &+L_{\mathrm{eff}}(\eta+\alpha_p)(1-\alpha_pL_{\mathrm{eff}}/2)].\end{aligned} \tag{6.10}$$

Comparing Eqs. (6.5), (6.7), and (6.10) with Eqs. (5.7)–(5.9) of Section 5, we conclude that the statistics of the PDG vector remains the same in the presence of polarization scrambling provided P_{in} is replaced by $P_{\text{in}}d_p$ in all expressions for the moments of PDG. As a result, the mean and rms value of PDG are reduced by a factor of d_p. This is exactly what has been observed experimentally.[30,31]

The general conclusion that the main effect of polarization scrambling is to reduce the input pump power by a factor of d_p as far as PDG is concerned allows us to translate all the results of Section 5 with this simple change. In particular, the probability distributions of the PDG vector $\mathbf{\Delta}$ and the PDG value Δ follow Eqs. (5.10) and (5.13), respectively. When the effective fiber length L_{eff} is much larger than the PMD diffusion length L_d, $\langle\Delta\rangle$ and σ_Δ are still given by approximations (5.14) and (5.15) except that P_{in} is replaced by $P_{\text{in}}d_p$. Although we assumed $\vec{\Gamma}_0$ and $\Gamma(z_2 - z_1)$ to be separable and $\Gamma(z_2 - z_1)$ to be exponential in the preceding discussion, the qualitative behavior is expected to be the same for any form of the correlation matrix $\vec{\Gamma}(z_2 - z_1)$ as long as $l_c \ll L_{\text{eff}}$.

7. SUMMARY

In this paper we have developed a general vector theory for analyzing fiber-based Raman amplifiers. We have shown that PMD can induce large fluctuations in the amplified signal depending on the value of the PMD parameter, although it also reduces the polarization dependence of the average gain. The amplification factor is found to follow the Gaussian statistics if it is measured in decibels, but the amplified signal itself follows a log-normal distribution.

We introduced a PDG vector to describe the nature of PDG in Raman amplifiers. We found that the probability distribution of the PDG mimics a Gaussian distribution when PMD is relatively small but becomes Maxwellian when the effective fiber length is much larger than the PMD diffusion length. Based on this probability distribution, we were able to find an analytic form of the dependence of the mean and variance of PDG on the operating parameters of Raman amplifiers. Both of these quantities depend inversely on the PMD parameter as well as on the frequency difference between the signal and the pump. We applied our vector theory to the case in which pump polarization is scrambled randomly and found that the mean PDG depends on the DOP of the pump polarization linearly.

We used the vector theory to compare the forward- and backward-pumping schemes used for making Raman amplifiers from the standpoint of PMD effects. In general, the use of backward pumping provides superior performance because it reduces the PDG as well as signal fluctuations to negligible levels as long as the PMD parameter D_p exceeds a relatively small value of $0.01\,\text{ps}/\sqrt{\text{km}}$. If forward pumping must be employed for practical reasons, one should use fibers with D_p values larger than $0.1\,\text{ps}/\sqrt{\text{km}}$. Physically speaking, backward pumping works better since the Stokes vectors $\mathbf{P}$ and $\mathbf{S}$ rotate in opposite directions and produce such rapid variations in their relative orientation angle θ that the PMD effects are averaged over a distance of 0.2 m or so.

In long-haul fiber links, PMD is intentionally reduced to minimize its effects on pulse broadening. However, our theory shows that PMD induces large signal fluctuations when the effective fiber length is comparable with the PMD diffusion length. Discrete Raman amplifiers typically use a few kilometers long fiber for providing sufficient gain. In the case of forward pumping, large signal fluctuations (>50%) are predicted to occur for D_p values $\sim 0.02\,\text{ps}/\sqrt{\text{km}}$. In the case of backward pumping, such D_P values can be tolerated but large signal fluctuations will reappear for $D_p \sim 0.001\,\text{ps}/\sqrt{\text{km}}$. If SRS is used for distributed amplification, one should balance the effect of PMD-induced pulse broadening and PMD-induced signal fluctuations carefully.

APPENDIX A

In this section we derive Eqs. (2.25) and (2.26) after transforming Eqs. (2.20) and (2.21) into a rotating frame and averaging over fast PMD-induced rotations at the pump frequency. We follow the averaging procedure described in Ref. 33 but carry it out in Stokes space rather than using the Jones-matrix notation. The required transformation in the Stokes space is governed by matrix $\vec{R}$ obtained by solving Eq. (2.24). We transform the Stokes vectors $\mathbf{P}$ and $\mathbf{S}$ into a reference frame rotating with $\vec{R}$, i.e., we introduce new Stokes vector $\mathbf{P}'$ and $\mathbf{S}'$ using $\mathbf{P} = \vec{R}\mathbf{P}'$ and $\mathbf{S} = \vec{R}\mathbf{S}'$. Equations (2.20) and (2.21) take the following form in the new rotating frame:

$$\xi\frac{d\mathbf{P}'}{dz} = -\alpha_p\mathbf{P}' - \frac{\omega_p g_1}{2\omega_s}[(1+3\mu)S_0\mathbf{P}' + (1+\mu)P_0\mathbf{S}' - 2\mu P_0\vec{R}^{-1}\mathbf{S}_3] + \mathbf{W}_p'\times\mathbf{P}', \tag{A1}$$

$$\frac{d\mathbf{S}'}{dz} = -\alpha_s\mathbf{S}' + \frac{g_1}{2}[(1+3\mu)P_0\mathbf{S}' + (1+\mu)S_0\mathbf{P}' - 2\mu S_0\vec{R}^{-1}\mathbf{P}_3] + (\Omega\mathbf{B} + \mathbf{W}_s')\times\mathbf{S}', \tag{A2}$$

where $\Omega = \omega_i - \xi\omega_p$, $\mathbf{B} = \vec{R}^{-1}\boldsymbol{\beta}$, and

$$\mathbf{W}_p' = \frac{2}{3}[\gamma_{pp}\vec{R}^{-1}\mathbf{P}_3 + 2\gamma_{ps}(1+\delta_b)\vec{R}^{-1}\mathbf{S}_3 - \gamma_{ps}(2+\delta_a+\delta_b)\mathbf{S}'], \tag{A3}$$

$$\mathbf{W}_s' = \frac{2}{3}[\gamma_{ss}\vec{R}^{-1}\mathbf{S}_3 + 2\gamma_{sp}(1+\delta_b)\vec{R}^{-1}\mathbf{P}_3 - \gamma_{sp}(2+\delta_a+\delta_b)\mathbf{P}']. \tag{A4}$$

Note that P_0 and S_0 remain unchanged because $\vec{R}$ does not change the length of Stokes vectors.

The transformed equations cannot be used in their present form because $\mathbf{W}_p'$ and $\mathbf{W}_s'$ depend on the random rotation operator $\vec{R}$. However, the pump frequency ω_p is so large in practice that $\vec{R}^{-1}\mathbf{P}_3$ and $\vec{R}^{-1}\mathbf{S}_3$ cover the entire surface of the Poincaré sphere over a short fiber sec-

tion. For this reason, we average them over all orientations using the general form of the rotation operator. In spherical coordinates, $\vec{\vec{R}}$ can be written as

$$\vec{\vec{R}} = \begin{pmatrix} \cos\theta & -\sin\theta\cos\varphi_0 & \sin\theta\sin\varphi_0 \\ \sin\theta\cos\varphi & \cos\theta\cos\varphi_0\cos\varphi - \sin\varphi_0\sin\varphi & -\cos\theta\sin\varphi_0\cos\varphi - \cos\varphi_0\sin\varphi \\ \sin\theta\sin\varphi & \cos\theta\cos\varphi_0\sin\varphi + \sin\varphi_0\cos\varphi & \cos\varphi_0\cos\varphi - \cos\theta\sin\varphi_0\sin\varphi \end{pmatrix}, \quad \text{(A5)}$$

where ϕ, ϕ_0, and θ are three independent random variables. Further, ϕ and ϕ_0 vary uniformly over the range from 0 to 2π whereas $\cos\theta$ varies uniformly in the range from -1 to 1.

To perform the average, we first note that

$$\vec{\vec{R}}^{-1}\mathbf{S}_3 = \vec{\vec{R}}^{-1}\hat{e}_3(\hat{e}_3\cdot\mathbf{S}) = \vec{\vec{R}}^{-1}\hat{e}_3(\hat{e}_3\cdot\vec{\vec{R}})\mathbf{S}'. \quad \text{(A6)}$$

Since $\hat{e}_3$ is a column vector (0, 0, 1), the column vector $\vec{\vec{R}}^{-1}\hat{e}_3$ and the row vector $\hat{e}_3\cdot\vec{\vec{R}}$ both correspond to the third row of matrix $\vec{\vec{R}}$. We can now average over θ, ϕ, and ϕ_0 using $\langle\sin^2\theta\rangle = 2/3$, $\langle\cos^2\theta\rangle = 1/3$, $\langle\sin^2\varphi\rangle = \langle\cos^2\varphi\rangle = 1/2$, and similar relations for other combinations of cosine and sine functions. Repeating the same procedure for $\vec{\vec{R}}^{-1}\mathbf{P}_3$, we obtain the simple result

$$\langle\vec{\vec{R}}^{-1}\mathbf{S}_3\rangle = \frac{1}{3}\mathbf{S}', \quad \langle\vec{\vec{R}}^{-1}\mathbf{P}_3\rangle = \frac{1}{3}\mathbf{P}'. \quad \text{(A7)}$$

Using these averages in Eqs. (A1) and (A2) and dropping the primes for simplicity of notation, we obtain Eqs. (2.25) and (2.26).

APPENDIX B

Here we provide details of the derivation of the averaged equations (3.2)–(3.7) that were used to find the average and the variance of the output signal. Our method follows the technique discussed in Ref. 24. The basic idea is to introduce a new vector $d\mathbf{W} = \mathbf{b}dz$ over a fiber segment of length dz that is much shorter than the total fiber length L but still much longer than the birefringence correlation length. Owing to the delta function correlation of the birefringence vector $\mathbf{b}$ in Eq. (2.29), the vector $d\mathbf{W}$ is a Wiener process with the following properties[24]:

$$\langle d\mathbf{W}\rangle = 0, \quad \langle d\mathbf{W}d\mathbf{W}\rangle = \frac{1}{3}D_p^2\vec{\vec{I}}dz,$$

$$\langle d\mathbf{W}\cdot d\mathbf{W}\rangle = D_p^2 dz. \quad \text{(B1)}$$

All other higher-order moments of $d\mathbf{W}$ vanish.

As discussed in Ref. 24, all the differential equations involving $d\mathbf{W}$ should be interpreted in the Stratonovich sense in the limit in which $\mathbf{b}$ reduces to a delta-correlated process from a general Markov process. In this interpretation, any nonanticipating function $f(z)$ that appears in the integral $\int f(z)d\mathbf{W}$ over a short length segment z to $z + dz$ is evaluated in the middle of the segment at the point $z + dz/2$. In contrast, the Ito calculus evaluates it at the left boundary point of the integral. As the Ito integral $\int f(z)d\mathbf{W}$ does not depend on the future, the average value of this integral vanishes. However, this does not happen for the Stratonovich integral because $f(z + dz/2)$ depends not only on its history, but also on $d\mathbf{W}$ within the interval from z to $z + dz$.

Consider now the stochastic differential equations (2.31) and (2.32). Before we can calculate any moments, we need to convert them from the Stratonovich to the Ito form. The conversion process is described in Chap. 4 of Ref. 24 and we follow it here. To illustrate the main steps, we consider Eq. (2.32) without the drift term and write it as

$$d\mathbf{s} = \mathbf{s}(z + dz) - \mathbf{s}(z) = -\Omega_R d\mathbf{W}\times\mathbf{s}(z + dz/2) \quad \text{(B2)}$$

(Stratonovich sense). Note that $\mathbf{s}$ on the right-hand side is evaluated in the middle of the segment. We expand it in a Taylor series as

$$\mathbf{s}(z + dz/2) = \mathbf{s}(z) + \frac{dz}{2}\frac{d\mathbf{s}(z)}{dz} + \cdots = \mathbf{s}(z) - \frac{\Omega_R}{2}d\mathbf{W}\times\mathbf{s}(z) + \cdots. \quad \text{(B3)}$$

Substituting this expansion in Eq. (B2), we obtain the Ito version of this equation as

$$d\mathbf{s} = -\Omega_R d\mathbf{W}\times\mathbf{s}(z) + \frac{\Omega_R^2}{2}d\mathbf{W}\times[d\mathbf{W}\times\mathbf{s}(z)] + \cdots. \quad \text{(B4)}$$

We now average over the second- and higher-order terms using the vector identity

$$d\mathbf{W}\times(d\mathbf{W}\times\mathbf{s}) = d\mathbf{W}(d\mathbf{W}\cdot\mathbf{s}) - \mathbf{s}(d\mathbf{W}\cdot d\mathbf{W}) \quad \text{(B5)}$$

and obtain the following Ito equation in the sense of a mean-square limit[24]:

$$d\mathbf{s} = -\Omega_R d\mathbf{W}\times\mathbf{s}(z) - \frac{1}{3}D_p^2\Omega_R^2\mathbf{s}(z)dz \quad \text{(B6)}$$

(Ito sense). The net effect of conversion is the appearance of a new term in the original Eq. (B2) when it is converted to the Ito sense. Following this procedure, Eqs. (2.31) and (2.32) can be converted to the following Ito version of these equations:

$$ds_0 = \frac{g_R}{2}\mathbf{P}\cdot\mathbf{s}dz, \quad \text{(B7)}$$

$$d\mathbf{s} = \frac{g_R}{2}\mathbf{P}s_0 dz - \eta\mathbf{s}dz + \Omega_R\mathbf{s}\times d\mathbf{W}, \quad \text{(B8)}$$

where $\eta = D_p^2\Omega_R^2/3$ as defined in the text. If we average Eqs. (B7) and (B8) over d$\mathbf{W}$, the last term in Eq. (B8) disappears, resulting in two deterministic equations. As $\hat{p}$ is fixed, we can introduce the angle θ using $\mathbf{s}\cdot\hat{p} = s_0\cos\theta$ and obtain Eqs. (3.2) and (3.3).

To calculate the second-order moments of $\mathbf{s}$, one can use Eqs. (2.31) and (2.32) to find the Stratonovich differential equations governing s_0^2, $s_0\mathbf{s}$, and $\mathbf{ss}$ and then convert them to the Ito form using the procedure described above. An alternative method uses Eqs. (B7) and (B8) together with the Ito formula given in Ref. 24. The final equations become

$$\mathrm{d}s_0^2 = g_R\mathbf{P}\cdot(s_0\mathbf{s})\mathrm{d}z, \tag{B9}$$

$$\mathrm{d}(s_0\mathbf{s}) = \frac{g_R}{2}[\mathbf{P}\cdot(\mathbf{ss}) + \mathbf{P}s_0^2]\mathrm{d}z - \eta(s_0\mathbf{s})\mathrm{d}z + \Omega_R(s_0\mathbf{s})\times\mathrm{d}\mathbf{W}, \tag{B10}$$

$$\mathrm{d}(\mathbf{ss}) = \frac{g_R}{2}(\mathbf{P}s_0\mathbf{s} + s_0\mathbf{sP}) - 3\eta(\mathbf{ss})\mathrm{d}z + \eta s_0^2\overleftrightarrow{I}\mathrm{d}z + \Omega_R(\mathbf{ss})\times\mathrm{d}\mathbf{W} - \Omega_R\mathrm{d}\mathbf{W}\times(\mathbf{ss}). \tag{B11}$$

When we average over d$\mathbf{W}$, all the terms containing d$\mathbf{W}$ disappear and the three equations become deterministic. Making inner products $s_0\mathbf{s}\cdot\hat{p}$ and $(\mathbf{s}\cdot\hat{p})^2$ and rewriting them as $s_0^2\cos\theta$ and $s_0^2\cos^2\theta$, we finally obtain Eqs. (3.6)–(3.8).

APPENDIX C

Here we give the derivation of Eq. (5.1), following the technique discussed in Ref. 27. Consider the evolution of an arbitrary Jones vector through the linear equation

$$\frac{\mathrm{d}|A\rangle}{\mathrm{d}z} = M(z)|A\rangle \equiv [m_0(z) + \mathbf{m}(z)\cdot\boldsymbol{\sigma}]|A\rangle. \tag{C1}$$

In general, $M(z)$ is not unitary, and $m_0(z)$ and $\mathbf{m}(z)$ can be complex. Since Eq. (C1) is linear, we can introduce a transfer matrix $T(z)$ as $|A_{\mathrm{out}}\rangle = T(z)|A_{\mathrm{in}}\rangle$, where $|A_{\mathrm{in}}\rangle$ and $|A_{\mathrm{out}}\rangle$ are input and output optical fields. This matrix satisfies

$$\frac{\mathrm{d}T(z)}{\mathrm{d}z} = [m_0(z) + \mathbf{m}(z)\cdot\boldsymbol{\sigma}]T(z). \tag{C2}$$

In terms of transfer matrix T, the input and output powers are related as

$$\langle A|A\rangle_{\mathrm{in}} = \langle A|[T(z)T^{\dagger}(z)]^{-1}|A\rangle_{\mathrm{out}}. \tag{C3}$$

The Hermitian matrix $T(z)T^{\dagger}(z)$ can be expanded in terms of the Pauli matrices as $T(z)T^{\dagger}(z) \equiv t_0(z) + \mathbf{t}(z)\cdot\boldsymbol{\sigma}$. It thus evolves as

$$\frac{\mathrm{d}[T(z)T^{\dagger}(z)]}{\mathrm{d}z} = \frac{\mathrm{d}t_0(z)}{\mathrm{d}z} + \frac{\mathrm{d}\mathbf{t}(z)}{\mathrm{d}z}\cdot\boldsymbol{\sigma}. \tag{C4}$$

In place of $\mathbf{t}$ we introduce $\mathbf{u}$ through the transformation

$$\mathbf{t}(z) = \mathbf{u}(z)\exp\left\{\int_0^z [m_0(z') + m_0^*(z')]\mathrm{d}z'\right\}. \tag{C5}$$

The same relation holds between t_0 and u_0. The new quantities $u_0(z)$ and $\mathbf{u}(z)$ are found to evolve with z as

$$\frac{\mathrm{d}u_0}{\mathrm{d}z} = 2\mathbf{m}_r\cdot\mathbf{u}, \tag{C6}$$

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}z} = 2\mathbf{m}_r u_0 - 2\mathbf{m}_i\times\mathbf{u}, \tag{C7}$$

where $\mathbf{m}_r$ and $\mathbf{m}_i$ are the real and imaginary parts of $\mathbf{m}$, respectively. Equations (C6) and (C7) show that $u_0^2 - u^2 = 1$ for all z, and thus the operator $(TT^{\dagger})^{-1}$ can be written as

$$[T(z)T^{\dagger}(z)]^{-1} = [u_0(z) - \mathbf{u}(z)\cdot\boldsymbol{\sigma}] \times \exp\left\{-\int_0^z \mathrm{d}z'[m_0(z') + m_0^*(z')]\right\}. \tag{C8}$$

We now apply Eq. (C3) to the signal being amplified by a Raman amplifier of length L and replace A with A_s. Using the definition $S_0 = \langle A_s|A_s\rangle$, where $S_0(z)$ is the signal power at a distance z, we obtain

$$S_0(0) = S_0(L)[u_0(L) - \mathbf{u}(L)\cdot\hat{s}_{\mathrm{out}}] \times \exp\left\{-\int_0^L [m_0(z) + m_0^*(z)]\mathrm{d}z\right\}, \tag{C9}$$

where $\hat{s}_{\mathrm{out}}$ is the unit vector in the direction of $\mathbf{S}(L) = \langle A_s|\boldsymbol{\sigma}|A_s\rangle$. The amplifier gain (in decibels) can now be written as

$$G = a\ln\left[\frac{S_0(L)}{S_0(0)}\right] = a\int_0^L [m_0(z) + m_0^*(z)]\mathrm{d}z - a\ln[u_0(L) - \mathbf{u}(L)\cdot\hat{s}_{\mathrm{out}}]. \tag{C10}$$

Note that G takes its maximum value when $\hat{s}_{\mathrm{out}}$ is along $\mathbf{u}$. In contrast, G is minimum when the two vectors are antiparallel.

The PDG can now be found by use of $\Delta = G_{\max} - G_{\min}$ and introducing the PDG vector as

$$\boldsymbol{\Delta} \equiv \hat{u}\Delta = a\hat{u}\ln\left(\frac{u_0 + u}{u_0 - u}\right), \tag{C11}$$

where $\hat{u}$ is the unit vector in the direction of $\mathbf{u}$ and $u = |\mathbf{u}|$. From Eqs. (C6) and (C7), $\boldsymbol{\Delta}$ is found to satisfy

$$\frac{\mathrm{d}\boldsymbol{\Delta}}{\mathrm{d}z} = \frac{\mathrm{d}u_0}{\mathrm{d}z}\frac{\partial\boldsymbol{\Delta}}{\partial u_0} + \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}z}\cdot\nabla_u(\boldsymbol{\Delta}) = 2\Delta\coth\left(\frac{\Delta}{2a}\right)[m_r - (m_r\cdot\hat{\Delta})\hat{\Delta}] + 4a(m_r\cdot\hat{\Delta})\hat{\Delta} - 2\mathbf{m}_i\times\boldsymbol{\Delta}, \tag{C12}$$

where ∇_u operates on $\mathbf{u}$ and $\hat{\Delta} \equiv \hat{u}$ is the unit vector in the direction of $\boldsymbol{\Delta}$.

We now need to find only $\mathbf{m}_r$ and $\mathbf{m}_i$. If we replace A with A_s in Eq. (C1) and follow the steps outlined in this

Appendix, we find that Eqs. (2.31) and (2.32) take the form of Eqs. (C6) and (C7). Comparing these equations, we obtain $\mathbf{m}_r = g_R\mathbf{P}/4$ and $\mathbf{m}_i = \Omega_R\mathbf{b}/2$. Using the parameters in Eq. (C12), we obtain the final result given as Eq. (5.1).

APPENDIX D

Here we provide details on the derivation of Eqs. (5.5)–(5.7) and (6.7). As discussed in Appendix B, Eq. (5.3) can be converted into an Ito equation as

$$d\mathbf{\Delta} = ag_R\mathbf{P}dz - \eta\mathbf{\Delta}dz + \Omega_R\mathbf{\Delta}\times d\mathbf{W}. \tag{D1}$$

From Eq. (D1) we can calculate $d\Delta^2$ and $d(\mathbf{\Delta\Delta})$ in the Ito sense and find

$$d\Delta^2 = 2ag_R\mathbf{P}\cdot\mathbf{\Delta}dz, \tag{D2}$$

$$d(\mathbf{\Delta\Delta}) = ag_R(\mathbf{P\Delta} + \mathbf{\Delta P})dz - 3\eta(\mathbf{\Delta\Delta})dz + \eta\Delta^2\vec{I}dz + \Omega_R(\mathbf{\Delta\Delta})\times d\mathbf{W} - \Omega_R d\mathbf{W}\times(\mathbf{\Delta\Delta}). \tag{D3}$$

When we average Eqs. (D1)–(D3) over $d\mathbf{W}$, all the terms containing $d\mathbf{W}$ disappear, and we obtain three deterministic equations as

$$d\langle\mathbf{\Delta}\rangle = ag_R\mathbf{P}dz - \eta\langle\mathbf{\Delta}\rangle dz, \tag{D4}$$

$$d\langle\Delta^2\rangle = 2ag_R\mathbf{P}\cdot\langle\mathbf{\Delta}\rangle dz, \tag{D5}$$

$$d\langle\mathbf{\Delta\Delta}\rangle = ag_R(\mathbf{P}\langle\mathbf{\Delta}\rangle + \langle\mathbf{\Delta}\rangle\mathbf{P})dz - 3\eta\langle\mathbf{\Delta\Delta}\rangle dz + \eta\langle\Delta^2\rangle\vec{I}dz. \tag{D6}$$

Using the definition of the covariance matrix $\vec{C}$ and noting that

$$d\vec{C} = d\langle\mathbf{\Delta\Delta}\rangle - (d\langle\mathbf{\Delta}\rangle)\langle\mathbf{\Delta}\rangle - \langle\mathbf{\Delta}\rangle(d\langle\mathbf{\Delta}\rangle), \tag{D7}$$

we finally obtain Eqs. (5.5)–(5.7).

In the case of pump polarization scrambling, $\mathbf{P}(z) = P_0(z)\hat{p}(z)$, where $\hat{p}(z)$ is random along the fiber. The covariance matrix $\vec{C}$ is then redefined as the double average over both random variables. The differential of $\vec{C}$ is then given by

$$d\vec{C} = d\langle\langle\mathbf{\Delta\Delta}\rangle\rangle - (d\langle\langle\mathbf{\Delta}\rangle\rangle)\langle\langle\mathbf{\Delta}\rangle\rangle - \langle\langle\mathbf{\Delta}\rangle\rangle(d\langle\langle\mathbf{\Delta}\rangle\rangle). \tag{D8}$$

By averaging Eqs. (D4)–(D6) over the ensemble of pump polarization, we obtain

$$d\langle\langle\mathbf{\Delta}\rangle\rangle = ag_R\langle\mathbf{P}\rangle dz - \eta\langle\langle\mathbf{\Delta}\rangle\rangle dz, \tag{D9}$$

$$d\langle\langle\Delta^2\rangle\rangle = 2ag_R\langle\mathbf{P}\cdot\langle\mathbf{\Delta}\rangle_b\rangle_p dz, \tag{D10}$$

$$d\langle\langle\mathbf{\Delta\Delta}\rangle\rangle = ag_R(\langle\mathbf{P}\langle\mathbf{\Delta}\rangle_b\rangle_p + \langle\langle\mathbf{\Delta}\rangle_b\mathbf{P}\rangle_p)dz - 3\eta\langle\langle\mathbf{\Delta\Delta}\rangle\rangle dz + \eta\langle\langle\Delta^2\rangle\rangle\vec{I}dz. \tag{D11}$$

Substituting Eqs. (D9) and (D11) into Eq. (D8), we obtain Eq. (6.7).

APPENDIX E

Here we provide details of the effects of NPR induced by the pump when pump polarization is scrambled. We neglect other effects temporarily and consider Eq. (2.26) with only the XPM-induced NPR term. After making the transformation in Eq. (6.2), Eq. (2.26) becomes

$$d\mathbf{S} = -\varepsilon_{sp}P_0(z)\delta\hat{p}(z + dz/2)\times\mathbf{S}(z + dz/2)dz \tag{E1}$$

(Stratonovich sense), where $\delta\hat{p}(z + dz/2) = \hat{p}(z + dz/2) - \langle\hat{p}\rangle$ and we have dropped the prime notation for simplicity. Since correlation length l_c is much smaller than the beat length of NPR (~10 km), we cut the fiber into many sections of l_c long, i.e., $dz = l_c$. According to the correlation matrix in Eq. (6.1), $\hat{p}(z)$ is approximately correlated within each section of length dz but is approximately uncorrelated from one section to another. Expanding $\mathbf{S}(z + dz/2)$ in Eq. (E1) into a Taylor series, this equation becomes

$$d\mathbf{S} = -\varepsilon_{sp}P_0(z)\delta\hat{p}(z + dz/2)\times\left[\mathbf{S}(z) + \frac{dz}{2}\frac{d\mathbf{S}}{dz} + \cdots\right]dz. \tag{E2}$$

Substituting Eq. (E1) into Eq. (E2), we obtain the Ito version of this equation as

$$d\mathbf{S} = -\varepsilon_{sp}P_0(z)\delta\hat{p}(z + dz/2)\times\mathbf{S}(z)dz + \frac{1}{2}[\varepsilon_{sp}P_0(z)dz]^2\{\delta\hat{p}(z)\delta\hat{p}(z + dz/2)\cdot\mathbf{S}(z) - [\delta\hat{p}(z)\cdot\delta\hat{p}(z + dz/2)]\mathbf{S}(z)\} + \cdots \tag{E3}$$

When we average this equation over $\hat{p}$, the first term disappears. As $\hat{p}$ is correlated within dz and does not change much, we can replace $\hat{p}(z + dz/2)$ with $\hat{p}(z)$ and perform the average. The final result is

$$d\langle\mathbf{S}\rangle \approx \frac{1}{2}[\varepsilon_{sp}P_0(z)dz]^2[\vec{\Gamma}_0 - \mathrm{Tr}(\vec{\Gamma}_0)\vec{I}]\cdot\langle\mathbf{S}(z)\rangle + \cdots = O[(dz)^2] \tag{E4}$$

in the sense of a mean-square limit.[24] When correlation length l_c is relatively small, $d\langle\mathbf{S}\rangle/dz \to 0$. As a result, NPR induced through pump polarization fluctuations does not diffuse the signal SOP. The net XPM-induced NPR contribution comes from only the average value $\langle\hat{p}\rangle$, as shown in Eq. (6.2). This result is easy to understand physically. It takes some distance for the signal to experience NPR. However, the pump polarization fluctuates so fast that the signal can only follow the rotation around the average pump polarization $\langle\hat{p}\rangle$.

ACKNOWLEDGMENT

This research is supported by the National Science Foundation under grants ECS-9903580 and DMS-0073923.

Q. Lin can be reached by e-mail at linq @optics.rochester.edu.

REFERENCES

1. C. V. Raman, "A new radiation," Indian J. Phys. **2**, 387–398 (1928).
2. E. J. Woodbury and W. K. Ng, "Ruby laser operation in the near IR," Proc. IRE **50**, 2347 (1962).
3. R. H. Stolen, E. P. Ippen, and A. R. Tynes, "Raman oscillation in glass optical waveguide," Appl. Phys. Lett. **20**, 62–64 (1972).

4. M. Ikeda, "Stimulated Raman amplification characteristics in long span single-mode silica fibers," Opt. Commun. **39**, 148–152 (1981).
5. S. Namiki and Y. Emori, "Ultrabroad-band Raman amplifiers pumped and gain-equalized by wavelength-division-multiplexed high-power laser diodes," IEEE J. Sel. Top. Quantum Electron. **7**, 3–16 (2001).
6. G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. (Academic, San Diego, Calif., 2001).
7. K. Rottwitt and A. J. Stentz, "Raman amplifiers in lightwave communication systems," in *Optical Fiber Telecommunications IV-A: Components*, I. P. Kaminow and T. Li, eds. (Academic, San Diego, Calif., 2002), Chap. 5.
8. R. H. Stolen, "Polarization effects in fiber Raman and Brillouin lasers," IEEE J. Quantum Electron. **QE-15**, 1157–1159 (1979).
9. D. J. Dougherty, F. X. Kartner, H. A. Haus, and E. P. Ippen, "Measurement of the Raman gain spectrum of optical fibers," Opt. Lett. **20**, 31–33 (1995).
10. D. Mahgerefteh, H. Yu, D. L. Butler, J. Goldhar, D. Wang, E. Golovchenko, A. N. Phlipetskii, C. R. Menyuk, and L. Joneckis, "Effect of randomly varying birefringence on the Raman gain in optical fibers," in *Conference on Lasers and Electro-Optics*, Vol. 11 of 1997 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1997), p. 447.
11. A. Berntson, S. Popov, E. Vanin, G. Jacobsen, and J. Karlsson, "Polarization dependence and gain tilt of Raman amplifiers for WDM systems," in *Optical Fiber Communication Conference*, Vol. 54 of OSA Trends in Optics and Photonics Series (Optical Society of America, Washington, D.C., 2001), pages MI2-1.
12. P. Ebrahimi, M. C. Hauer, Q. Yu, R. Khosravani, D. Gurkan, D. W. Kim, D. W. Lee, and A. E. Willner, "Statistics of polarization dependent gain in Raman fiber amplifiers due to PMD," in *Conference on Lasers and Electro-Optics*, Vol. 56 of OSA Trends in Optics and Photonics Series (Optical Society of America, Washington, D.C., 2001), pp. 143–144.
13. S. Popov, E. Vanin, and G. Jacobsen, "Influence of polarization mode dispersion value in dispersion-compensating fibers on the polarization dependence of Raman gain," Opt. Lett. **27**, 848–850 (2002).
14. J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization mode dispersion in optical fibers," Proc. Natl. Acad. Sci. USA **97**, 4541–4550 (2000).
15. S. J. Savory and F. P. Payne, "Pulse propagation in fibers with polarization-mode dispersion," J. Lightwave Technol. **19**, 350–357 (2001).
16. D. Wang and C. R. Menyuk, "Calculation of penalties due to polarization effects in a long-haul WDM system using a Stokes parameter model," J. Lightwave Technol. **19**, 487–494 (2001).
17. H. Kogelnik, R. M. Jopson, and L. E. Nelson, "Polarization-mode dispersion," in *Optical Fiber Telecommunications IV-B: Systems and Impairments*, I. P. Kaminow and T. Li, eds. (Academic, San Diego, Calif., 2002), Chap. 15.
18. Q. Lin and G. P. Agrawal, "Polarization mode dispersion-induced fluctuations during Raman amplification in optical fibers," Opt. Lett. **27**, 2194–2196 (2002).
19. Q. Lin and G. P. Agrawal, "Statistics of polarization-dependent gain in fiber-based Raman amplifiers," Opt. Lett. **27**, 227–229 (2003).
20. R. Hellwarth, J. Cherlow, and T. Yang, "Origin and frequency dependence of nonlinear optical susceptibilities of glasses," Phys. Rev. B **11**, 964–967 (1975).
21. R. W. Hellwarth, "Third-order optical susceptibilities of liquid and solids," Prog. Quantum Electron. **5**, 1–68 (1977).
22. P. Ciprut, B. Gisin, N. Gisin, R. Passy, J. P. Von der Weid, F. Prieto, and C. W. Zimmer, "Second-order polarization mode dispersion: impact on analog and digital transmissions," J. Lightwave Technol. **16**, 757–771 (1998).
23. R. W. Boyd, *Nonlinear Optics*, 2nd ed. (Academic, San Diego, Calif., 1992).
24. C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer-Verlag, New York, 1985).
25. H. H. Kee, C. R. S. Fludger, and V. Handerek, "Statistical properties of polarization dependent gain in fiber Raman amplifiers," in *Optical Fiber Communication*, Vol. 70 of 2002 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 2002), pp. 180–181.
26. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. (WCB/McGraw-Hill, New York, 1991).
27. B. Huttner, C. Geiser, and N. Gisin, "Polarization-induced distortions in optical fiber networks with polarization-mode dispersion and polarization-dependent losses," IEEE J. Sel. Top. Quantum Electron. **6**, 317–329 (2000).
28. A. Mecozzi and M. Shtaif, "The statistics of polarization-dependent loss in optical communication systems," IEEE Photon. Technol. Lett. **14**, 313–315 (2002).
29. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," J. Lightwave Technol. **9**, 1439–1456 (1991).
30. Y. Emori, S. Matsushita, and S. Namiki, "Cost-effective depolarized diode pump unit designed for C-band flat-gain Raman amplifiers to control EDFA gain profile," in *Optical Fiber Communication Conference*, Vol. 37 of OSA Trends in Optics and Photonics Series (Optical Society of America, Washington, D.C., 2000), pp. 106–107.
31. J. Zhang, V. Dominic, M. Missey, S. Sanders, and D. Mehuys, "Dependence of Raman polarization dependent gain on pump degree of polarization at high gain levels," in *Optical Amplifiers and Their Applications*, A. Mecozzi, M. Shimizu, and J. Zyskind, eds., Vol. 44 of OSA Trends in Optics and Photonics (Optical Society of America, Washington, D.C., 2000), pp. 30–31.
32. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University, New York, 1995), Chap. 6.
33. P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," J. Lightwave Technol. **14**, 148–157 (1996).