# Parametric-Gain Approach to the Analysis of Single-Channel DPSK/DQPSK Systems With Nonlinear Phase Noise

Paolo Serena, Alessandra Orlandini, and Alberto Bononi

*Abstract*—This paper presents a novel method based on a parametric gain (PG) approach to study the impact of nonlinear phase noise in single-channel dispersion-managed differentially phase-modulated systems. This paper first shows through Monte Carlo simulations that the received amplified spontaneous emission (ASE) noise statistics, before photodetection, can be reasonably assumed to be Gaussian, provided a sufficiently large chromatic dispersion is present in the transmission fiber. This paper then evaluates in a closed form the ASE power spectral density by linearizing the interaction between a signal and a noise in the limit of a distributed system. Even if the received ASE is nonstationary in time due to pulse shape and modulation, this paper shows that it can be approximated by an equivalent stationary process, as if the signal were continuous wave (CW). This paper then applies the CW-equivalent ASE model to bit-error-rate evaluation by using an extension of a known Karhunen–Loéve method for quadratic detectors in colored Gaussian noise. Such a method avoids calculation of the nonlinear phase statistics and accounts for intersymbol interference due to a nonlinear waveform distortion and optical and electrical postdetection filtering. This paper compares binary and quaternary schemes with both nonreturn- and return-to-zero (RZ) pulses for various values of nonlinear phases and bit rates. The results confirm that PG deeply affects the system performance, especially with RZ pulses and with quaternary schemes. This paper also compares ON–OFF keying (OOK) differential phase-shifted keying (DPSK) systems, showing that the initial 3-dB advantage of DPSK is lost for increasing nonlinear phases because DPSK is less robust to PG than OOK.

*Index Terms*—Differential phase-shift keying (DPSK), differential quadrature phase-shift keying (DQPSK), Karhunen–Loéve (KL) transforms, nonlinear phase noise, parametric gain (PG).

## I. INTRODUCTION

OPTICAL phase-shift keying (PSK) modulation formats are a promising technique to increase the performance of long-haul transmission systems [1]–[4]. Compared with the conventional ON–OFF keying (OOK) format, PSK formats detected with an optical delay demodulator and a balanced receiver have the major benefit of a lower optical signal-to-noise ratio (OSNR) requirement, which leads to an increased system margin and an extended transmission distance, or equivalently to reduced a transmitted power and thus an increased tolerance to fiber nonlinearities.

However, unlike OOK, the benefits of PSK formats may be severely limited by nonlinear phase noise [5], i.e., the amplitude-to-phase-noise conversion due to the nonlinear interaction of signal and amplified spontaneous emission (ASE) noise during propagation. Such an interaction manifests itself also as a parametric gain (PG) of the received ASE noise. PG changes the statistics and the correlation of the ASE noise, which does not remain white during propagation. Moreover, in presence of PG and return-to-zero (RZ) pulses whose intensity is not constant, the ASE statistics are time dependent, and thus the ASE is a nonstationary process.

Attempts have been made to theoretically study the statistics of the nonlinear phase noise in order to assess the bit error rate (BER) of binary differential PSK (DPSK) receivers based on ideal phase discriminators, but an exact expression of these statistics was found only at zero group-velocity dispersion (GVD) [6], [7] while an approximate expression including GVD based on a phenomenological argument was presented in [8].

In this paper, we take a different approach to BER evaluation, which does not need the phase noise statistics. We first numerically show that in practical dispersion-managed (DM) systems, working at a sufficiently large OSNR and in which some local GVD is present, the received ASE noise before photodetection can still be reasonably modeled by a Gaussian process. Thanks to this result, the system BER can be evaluated through a rigorous Karhunen–Loéve (KL) method for square-law detectors, which leads a generalized chi-square distribution of the sampled current in the electrical domain [9]. Then, focusing on a practical balanced PSK receivers based on Mach–Zehnder delay demodulators, we extend the KL method to DPSK/DQPSK modulation formats, much like what is done in [9] and [10]. The main contribution of our paper is in the extension of the method to the case of a nonwhite ASE noise before demodulation, providing a BER that accounts not only for the intersymbol interference (ISI) due to propagation, optical, and electrical postdetection filtering, but also for noise PG. All the analytical details needed for the implementation of the extended KL method are provided in the paper. Moreover, the KL expansion over the Fourier basis [9] allows a more straightforward numerical implementation of the BER algorithm, as compared with an alternative method that explicitly computes the KL eigenfunctions through the solution of an integral equation [11].

For BER computation with PG, under the above assumption of Gaussian optical ASE, the exact nonstationary power
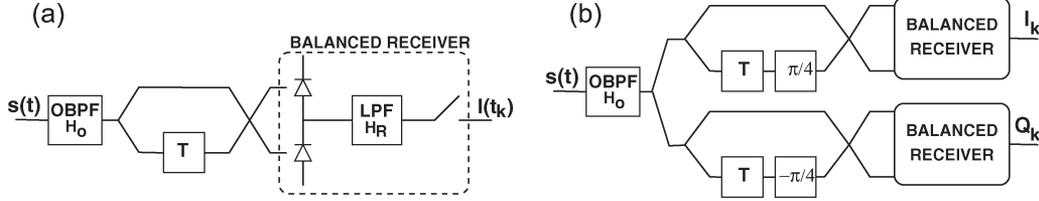
Fig. 1.  Scheme of the (a) DPSK and (b) DQPSK receiver.
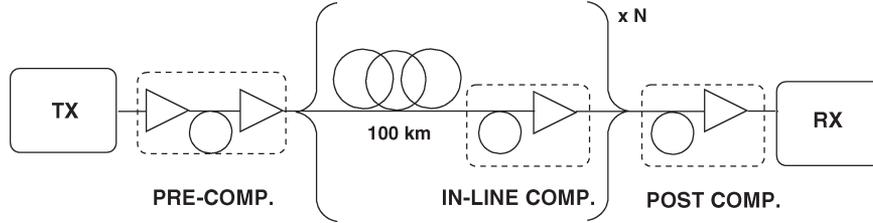


Fig. 2.  Setup of the single-channel DM DPSK/DQPSK system.

spectral density (PSD) of the ASE noise can be obtained through a computationally intensive method based on a linearization of the nonlinear Schrödinger equation (NLSE) [12] around the noiseless solution obtained through the standard split-step Fourier method (SSFM). However, its computational complexity grows as the cube of the number of ASE time samples, which makes the method extremely time consuming.

In this paper, we get around such an inefficiency by proposing and verifying a novel model, based also on a linearized version of the NLSE, which leads to an *equivalent stationary* ASE process obtained by using a continuous wave (CW) signal of properly modified peak power, which we call the CW-equivalent ASE process.

We then apply our novel KL algorithm for BER evaluation to a comparative study of the performance of single-channel DM DPSK and differential quadrature PSK (DQPSK) systems, both for nonreturn-to-zero (NRZ) and RZ supporting pulses. Results are provided in terms of OSNR penalties for different DM system parameters.

The paper is organized as follows. Section II describes the PSK system setup studied throughout the paper. In Section III, the Gaussian assumption for the received ASE noise is discussed in DM terrestrial systems of practical interest. Section IV is devoted to the small-signal model description and its application to the ASE PSD computation, both in the NRZ and RZ case. Section V describes the algorithm for BER evaluation for DPSK/DQPSK signals, both with and without PG. In Section VI, we numerically verify the accuracy of our model for BER computation and we provide performance of comparisons of DPSK/DQPSK systems, as well as OOK/DPSK systems. In Section VII, we draw our main conclusions.

## II. PSK MODULATION FORMATS

Two commonly used implementations of optical PSK formats are based on differential binary and quaternary schemes [2], [3]. DPSK modulation format is a binary modulation which encodes the information onto the differential optical phase $\Delta\Phi$

between adjacent bits, which can be either 0 or $\pi$. The typical transmitter for NRZ-DPSK requires a CW laser source followed by a phase modulator or, alternatively, a Mach–Zehnder modulator, which is driven by the information data to be encoded and generates an optical phase of 0, $\pi$. In the case of RZ-DPSK, a second Mach–Zehnder modulator, synchronized to the data, is used to carve the output RZ pulses with the required duty cycle. The DQPSK format utilizes two-orthogonal DPSK signals, which are differentially encoded with two independent data streams and then combined. The signal phase on each symbol belongs to the alphabet $\{\pm(\pi/4), \pm(3\pi/4)\}$, which leads to the four possible values 0, $\pi/2$, $\pi$, $3/2\,\pi$ of the differential phase $\Delta\Phi$. Since for DQPSK, each value of the optical differential phase corresponds to a pair of encoded bits, the aggregate signal bit rate, and thus its spectral efficiency, is doubled with respect to DPSK.

As shown in Fig. 1(a), the detection of the DPSK signal is commonly based on a Mach–Zehnder interferometer with a delay equal to the bit time $T$, followed by the balanced receiver where half the sum of the input fields at times $t$ and $t - T$ is detected by one photodetector, while half the difference is detected by the other [13]. The DQPSK receiver [see Fig. 1(b)] is usually based on two DPSK demodulators, which independently detect the in-phase and quadrature signal components with an additional phase shift of, respectively, $\pi/4$ and $-\pi/4$ in the Mach–Zehnder interferometer [2].

In Fig. 2, we show the scheme of the single-channel DPSK/DQPSK multispan DM system that will be studied in the following sections. All spans, each composed of a 100-km-long transmission fiber followed by a dispersion-compensating fiber, have a zero in-line residual dispersion (full span compensation). For each launched power value, pre- and postcompensating fibers, placed before and after the transmission line, are always optimized in order to minimize the BER. The optical filter at the receiver is Gaussian with bandwidth $B_o = 1.8R$, being $R$ the system bit rate. After the balanced receiver, the difference between the received currents is filtered by a Bessel fifth-order filter of bandwidth $B_e = 0.65R$ and then sampled.
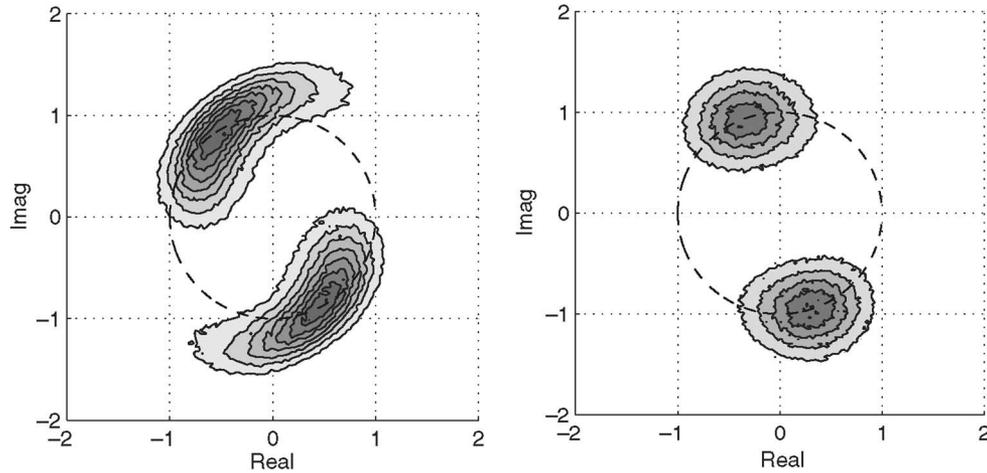
Fig. 3. Contour plots of ASE joint pdf on an NRZ-DPSK signal before the receiver for OSNR = 25 dB/0.1 nm. Fully compensated span with $D_{\text{tx}} = 0$ ps/nm/km (left) and $D_{\text{tx}} = 4$ ps/nm/km (right). $\overline{\Phi}_{\text{NL}} = 0.3\,\pi$rad.

### III. RECEIVED ASE STATISTICS

The ASE noise and the transmitted signal interact during propagation through a four-wave mixing process that colors the PSD of the initially white ASE noise components, both in-phase and in-quadrature with the signal [14]. It is known that signal and ASE noise have maximum nonlinear interaction strength at zero GVD, yielding ASE statistics that strongly depart from Gaussian [6], [7]. In this section, we numerically show that, in a more practical DM system in which the transmission fiber has nonzero GVD for effective four-wave mixing suppression, the joint probability density function (pdf) of the in-phase and quadrature-received ASE components before detection can be reasonably approximated with a Gaussian distribution. To this aim, we estimated the joint pdf of the received ASE before optical filtering by a Monte Carlo simulation ($2^{19}$ time samples) of an NRZ-DPSK signal corrupted by the preamplifier noise at the transmitter. The received OSNR in back to back was 25 dB over the conventional bandwidth of 0.1 nm. The transmission link was a single fully compensated span with transmission fiber chromatic dispersion $D_{\text{tx}}$. The average cumulated nonlinear phase was $\overline{\Phi}_{\text{NL}} = 0.3\,\pi$ rad, according to the definition

$$\overline{\Phi}_{\text{NL}} = \gamma \overline{P} L_{\text{eff}} N \tag{1}$$

where $\gamma$ is the nonlinear fiber coefficient, $\overline{P}$ indicates the time-averaged launched power, $L_{\text{eff}} = \int_0^L e^{-\alpha(\xi)} d\xi$ is the effective length of the transmission fiber with length $L$ and attenuation $\alpha$, $N$ is the number of spans, equal to 1 in this case. In (1), we neglected any nonlinearity in the compensating fibers. In Fig. 3, we plot the contour levels of the joint pdf of the imaginary (quadrature) and real (in-phase) optical ASE components, for $D_{\text{tx}} = 0$ (left) and $D_{\text{tx}} = 4$ ps/nm/km (right), down to a value of $2 \times 10^{-6}$. The left figure confirms that Kerr nonlinearity alone reshapes the input Gaussian ASE noise into a non-Gaussian one. On the contrary, with a nonzero transmission fiber GVD, we obtained the contours shown in Fig. 3 (right), whose elliptical shapes match those of a Gaussian bivariate distribution, at least down to the precision of the Monte Carlo simulation. Being the single span, a worst case for the Gaussian

assumption, such a result holds true also for multiple spans at the same overall nonlinear phase rotation. In fact, the increased number of independent ASE sources provided by the in-line amplifiers greatly accelerates the convergence to a Gaussian pdf, according to the central limit theorem [15]. Note however that the Gaussian assumption becomes unreliable at either very low-OSNR values or at very large nonlinear phases. In Section IV, we will make this statement more precise by providing a simple formula, based on the results in [16], relating the largest nonlinear phase for the Gaussian assumption to hold and the key DM system parameters.

### IV. MODEL FOR ASE PROPAGATION

In this section, we calculate the ASE noise statistics starting from the NLSE of a single-channel periodic DM system. The electric field $A(z, t)$, where $z$ is the distance and $t$ the time normalized to the mark duration $d \cdot T$, where $d$ is the duty cycle, propagates in its retarded time frame as

$$\frac{\partial A}{\partial z} = j \frac{1}{2 L_{\text{d}}(z)} \frac{\partial^2 A}{\partial t^2} - j \frac{1}{L_{\text{NL}}(z)} |A|^2 A + \frac{g(z)}{2} A \tag{2}$$

where $A(z, t)$ is normalized to the square root of the transmitted peak power $P_{\text{peak}}$, $L_{\text{NL}}(z) = 1/\gamma(z) P_{\text{peak}}$ is the local nonlinear length; $L_{\text{d}}(z) = (d \cdot T)^2 / \beta_2(z)$ is the local dispersion length referred to the mark duration; $g(z) = -(1/L_A(z)) + \Sigma_k G_k \delta(z - kL)$ is the net logarithmic gain/attenuation per unit length, where $L_A(z) = 1/\alpha(z)$ is the fiber attenuation length, and $e^{G_k}$ is the power gain of the $k$th lumped amplifier of the link placed at $z = kL$. $\delta(.)$ indicates the Dirac's delta function. $L_{\text{d}}(z)$, $L_{\text{NL}}(z)$, $L_A(z)$ and $g(z)$ are $z$-periodic functions with period equal to the span length $L$.

We now normalize $A(z, t)$ to the net fiber gain/attenuation up to $z$, i.e.,

$$A(z, t) = U(z, t) e^{(1/2) \int_0^z g(x) dx} = U(z, t) \sqrt{f(z)}$$

being $f(z) = \exp(\int_0^z g(x)dx)$. Thus, (2) becomes

$$\frac{\partial U}{\partial z} = j\frac{1}{2L_{\mathrm{d}}(z)}\frac{\partial^2 U}{\partial t^2} - j\frac{f(z)}{L_{\mathrm{NL}}(z)}|U|^2 U + W_A(z,t) \quad (3)$$

where we have also included a zero-mean Gaussian noise term $W_A(z,t)$ with autocorrelation $R(z_1, z_2, t_1, t_2) = E\{W_A(z_1,t_1)W_A^*(z_2,t_2)\}$ at times $t_1, t_2$ and coordinates $z_1, z_2$ equal to

$$R(z_1, z_2, t_1, t_2) = \delta(t_1 - t_2)\delta(z_1 - z_2)\sum_k N_{0k}\delta(z_1 - kL)$$

where $N_{0k}$ is the white one-sided PSD of each ASE polarization introduced by the $k$th amplifier, the asterisk denotes complex conjugate, and $E\{.\}$ indicates statistical averaging.

If the transmitted field is a CW, in absence of noise the solution of (3) is $U(z) = e^{-j\Phi_{\mathrm{NL}}(z)}$, where $\Phi_{\mathrm{NL}}(z) = \int_0^z f(x)(1/L_{\mathrm{NL}}(x))dx$ is the nonlinear phase cumulated by the CW, in agreement with (1). By adding the noise contribution, we search for a perturbed solution of (3) of the kind [14], [17]

$$U(z,t) = (1 + u(z,t))\,e^{-j\Phi_{\mathrm{NL}}(z)} \quad (4)$$

where $u(z,t)$ accounts for the noise. By inserting (4) into (3), and by assuming $|u|^2 \ll 1$, so that higher order powers of $u(z,t)$ can be dropped [12], [14], we obtain the linearized NLSE for the perturbation

$$\frac{\partial u}{\partial z} = j\frac{1}{2L_{\mathrm{d}}(z)}\frac{\partial^2 u}{\partial t^2} - j\frac{f(z)}{L_{\mathrm{NL}}(z)}(u + u^*) + W_A \quad (5)$$

where the phase rotation $\Phi_{\mathrm{NL}}$ in (4) has not changed the statistics of $W_A$. By indicating the Fourier transforms of $u(z,t)$ and $W_A(z,t)$, respectively, with $\tilde{u}(z,\omega)$ and $\widetilde{W}_A(z,\omega)$, where $\omega$ is the angular frequency normalized to $R/d$, (5) in the frequency domain rewrites as

$$\frac{\partial \tilde{u}}{\partial z} = -j\frac{\omega^2}{2L_{\mathrm{d}}(z)}\tilde{u}(z,\omega)$$
$$- j\frac{f(z)}{L_{\mathrm{NL}}(z)}\left[\tilde{u}(z,\omega) + \tilde{u}^*(z,-\omega)\right] + \widetilde{W}_A(z,\omega). \quad (6)$$

Thanks to its $z$-periodic behavior, the local dispersion length can be written as $1/L_{\mathrm{d}}(z) = (1/L_{\mathrm{d}}) + (1/L_{\Delta}(z))$, where $1/L_{\mathrm{d}} = (1/L)\int_0^L (1/L_{\mathrm{d}}(x))dx$ is the inverse span-averaged dispersion length, while $1/L_{\Delta}$ accounts for the local deviation from such an average. Inside each span, we recognize two different dynamics along $z$ due to the fiber dispersion, a *slow* dynamic due to $L_{\mathrm{d}}$ and a *fast* dynamic due to $L_{\Delta}$. We next move into a reference system that follows the fast dynamic by making the change of variable

$$\tilde{u}(z,\omega) = \tilde{a}(z,\omega)e^{-j\frac{\Theta_\Delta(z,\omega)}{2}} \quad (7)$$

where $\Theta_\Delta = \omega^2 \int_0^z (1/L_{\Delta}(x))dx$. Substituting (7) in (6) yields

$$\frac{\partial \tilde{a}}{\partial z} = -j\frac{\omega^2}{2L_{\mathrm{d}}}\tilde{a}(z,\omega) - j\frac{f(z)}{L_{\mathrm{NL}}(z)}$$
$$\cdot\left[\tilde{a}(z,\omega) + \tilde{a}^*(z,-\omega)e^{j\Theta_\Delta}\right] + \widetilde{W}_A(z,\omega) \quad (8)$$

where again, the phase rotation in (7) does not change the Gaussian statistics of the noise $\widetilde{W}_A$. For a finite received nonlinear phase, when the number of spans $N \to \infty$, the infinitesimal nonlinear phase rotation per span turns out to drive the evolution of $a(z,\omega)$ as a slowly varying $z$-function span by span. Thus, drawing upon the idea of separating the fast and slow noise dynamics, $a(z,\omega)$ cannot follow the fast variations within each span due to $L_{\mathrm{NL}}(z)$, $f(z)$, and $\Theta_\Delta(z)$, but only their average effect. Hence, using the method of averaging [18], we substitute the rapidly varying terms in (8) with their span-averaged values

$$\left\langle \frac{f(z)e^{j\Theta_\Delta(z,\omega)}}{L_{\mathrm{NL}}(z)} \right\rangle = \frac{1}{L}\int_0^L \frac{f(x)e^{j\Theta_\Delta(x,\omega)}}{L_{\mathrm{NL}}(x)}dx \stackrel{\triangle}{=} \mathcal{R}(\omega). \quad (9)$$

Introducing the kernel of the transmission link $r(\omega) \stackrel{\triangle}{=} \mathcal{R}(\omega)/\mathcal{R}(0)$ [19], (8) rewrites as

$$\frac{\partial \tilde{a}}{\partial z} = -j\frac{\omega^2}{2L_{\mathrm{d}}}\tilde{a}(z,\omega) - j\mathcal{R}(0)$$
$$\cdot\left[\tilde{a}(z,\omega) + \tilde{a}^*(z,-\omega)r(\omega)\right] + \widetilde{W}(z,\omega) \quad (10)$$

where, thanks to the method of averaging, we substituted the white ASE $\widetilde{W}_A$ with a Langevin–Gaussian noise process $\widetilde{W}$ [20] with PSD at coordinates $(z_1, z_2)$ equal to $E\{\widetilde{W}(z_1,\omega)\widetilde{W}^*(z_2,\omega)\} = 2\sigma^2\delta(z_1 - z_2)$, being $2\sigma^2$ the one-sided ASE PSD per unit length. For an $N$-span link it is $2\sigma^2 NL = \sum_{k=1}^N N_{0k}$. In terrestrial systems having long spans ($L \gg L_A$), we find

$$r(\omega) \cong \frac{1}{1 + jS\omega^2} \quad (11)$$

where we call $S \stackrel{\triangle}{=} -(L_A/L_\Delta)$ the *map strength* of the terrestrial DM system[1] [21], [22]. It is also $\mathcal{R}(0) = (L_A/L) \cdot (1/L_{\mathrm{NL}})$, which corresponds to the inverse span-averaged nonlinear length. Note that all parameters in the system lengths $L_A$, $L_{\mathrm{NL}}$, $L_\Delta$ refer to the transmission fiber.

In the Appendix, we discuss the analytical details of the solution of (10). Such a solution yields a stationary Gaussian noise $\tilde{a}(z,\omega)$ whose in-phase and quadrature components are correlated by PG and whose normalized PSDs have the closed-form expression given in (28). Such an expression can be shown to coincide with the PSDs obtained in [23] in the limit $N \to \infty$.

It is easy to verify that (10) is a linearization of the DM-NLSE of Ablowitz and Hirooka [19]. The limits of applicability of the method of averaging are discussed at length in [24], where it is shown that the method holds for small nonlinear phase rotations per span. We verified that the DM-NLSE yields very accurate PSDs when the nonlinear phase rotation per span is roughly below 0.02 rad and qualitatively reasonable results up to nonlinear phase rotations per span of 0.1 rad.

Closer examination of (28) reveals that, in the long-span terrestrial map case whose kernel $r(\omega)$ is given in (11), the

---

[1]Note that we associate to $S$ the sign of the transmission fiber dispersion $D_{\mathrm{tx}}$.

ASE PSD solely depends on the three following dimensionless parameters: 1) the strength $S$; 2) the normalized average in-line dispersion $z/L_d$; and 3) the peak nonlinear phase $\Phi_{NL}(z) = (z/L_{NL})(L_A/L)$. These three parameters provide general scaling rules both for the noise PG analysis and for the general design of DM terrestrial systems.

The model described by (10) is linear, so that ASE keeps its initially Gaussian statistics also after propagation. In [16], a simple bound on the average cumulated nonlinear phase was provided, below which, the Gaussian ASE assumption was guaranteed to hold in the evaluation of system penalties in long-haul DM terrestrial OOK NRZ systems. Such a bound [16, Eq. (11)], written in terms of our normalized system parameters, is

$$\overline{\Phi}_{NL} = \left( 12\sqrt{2\,|S|}\,\kappa^3 \frac{\Delta\nu}{R/d} \mathrm{OSNR} \right)^{\frac{1}{4}} \tag{12}$$

where

$$\mathrm{OSNR} = \frac{\overline{P}}{2N_0\Delta\nu} \tag{13}$$

is the OSNR in absence of PG, and where $\Delta\nu = 12.5$ GHz (0.1 nm) is the conventional bandwidth for the OSNR measurement, $\kappa = \overline{P}/P_{peak}$ is the average energy per bit, and $N_0 = 2\chi\sigma^2 z$ is the one-sided PSD for each ASE noise polarization. $\chi = (d/R)P_{peak}$ is a conversion factor from normalized to standard units.

### A. Effects of Pulse Shape on ASE

Since the proposed ASE model is based on a CW assumption for the signal, it does not take into account the influence of a signal modulation and pulse shape. We investigated the above subject by first measuring the PSD of the complex signal $a(z,t)$ for a DM NRZ-DPSK system, i.e., one with constant input power. We compared the exact PSD computed through the algorithm of Holzlöhner *et al.* [12] for a specific modulating bit sequence, and that of the proposed solution based on a CW signal of power equal to the normalized peak power $P_{peak} = 1$. According to our stationary ASE model, the normalized PSD of $a(z,t)$ coincides with the trace of the matrix $\mathbf{G}$ in (28). Since the true PSD provided by the numerical algorithm in [12] is nonstationary, for the comparison, we evaluated it at times $t_s$ placed at the center of the information bits. The analyzed $20 \times 100$ km DM system had $D_{tx} = 8$ ps/km/nm, and $z/L_d = 0$, with optimized pre- and postcompensating fibers. The average cumulated nonlinear phase was $\overline{\Phi}_{NL} = 0.3\pi$. Fig. 4 shows our CW-equivalent PSD (dashed line) and the exact PSDs (algorithm in [12], solid lines), one for each bit, versus normalized frequency $f_n = \omega/2\pi$, at a strength $S = 0.03$ (i.e., $R \sim 10$ Gb/s). The exact nonstationary PSD traces are found to almost coincide at all considered sampling times, so that the received ASE is quasi-stationary in time [22]. The good match with our CW-equivalent ASE PSD was expected, since here the normalized signal power remains almost constant in time and equal to one.

On the other hand, for RZ-DPSK systems, the received noise is nonstationary in time and its PSD follows the signal power
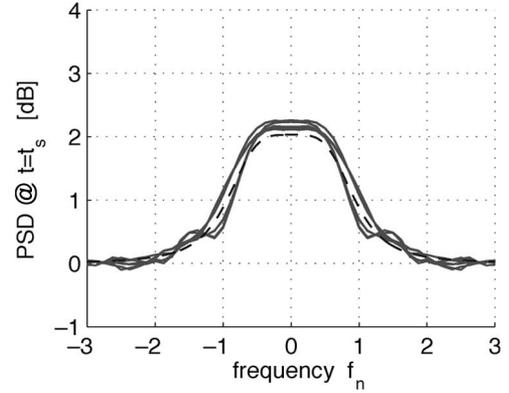


Fig. 4. Normalized ASE PSD versus frequency $f_n$ for a DM $20 \times 100$ km NRZ-DPSK system with $z/L_d = 0$, $\overline{\Phi}_{NL} = 0.3\pi$, $S = 0.03$. Dashed line: CW-equivalent ASE PSD (28); solid lines: exact PSDs at each sampling time [12].
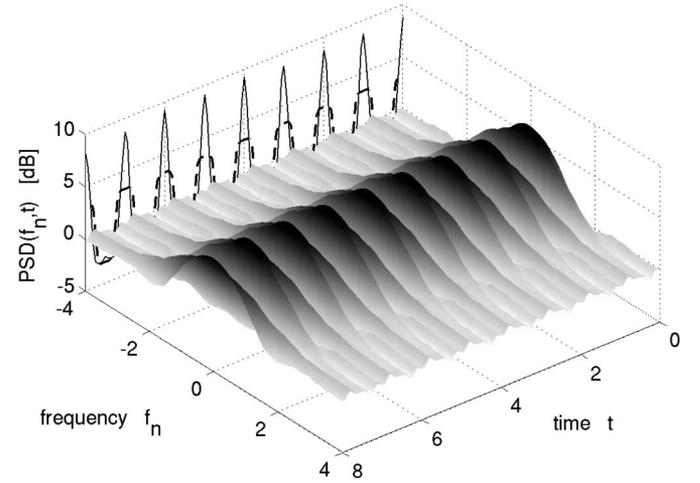


Fig. 5. Exact normalized ASE PSD [12] versus frequency $f_n$ and time $t$ for the same link of Fig. 4 but RZ-DPSK modulation and $S = 0.1$. On the back: transmitted (dashed line) and received (solid line) power in arbitrary units.

profile $P(t)$, as shown in Fig. 5 for a duty cycle $d = 0.5$. Here, we plot the exact PSD [12] versus frequency and time for the same link of Fig. 4 but at strength $S = 0.1$. In the figure, we also show the power (dashed line) of 8 of the 16 transmitted bits and the corresponding received signal power (solid line) in arbitrary units. In spite of the nonstationarity of the noise, the PSD is almost cyclostationary with period equal to the bit time, which suggests that an equivalent quasi-stationary model for the noise PSD should exist. We therefore expect that our CW-equivalent model still applies by using as the CW reference power a proper effective value $P_{eff}(t_s)$ at times $t_s$. The intuition about the appropriate value of $P_{eff}$ comes from (10), which reveals that, at a specific $\omega$, the noise field gets energy from the CW only within a proper frequency bandwidth. Such a bandwidth, which is essentially set by the bandwidth of the kernel $r(\omega)$, corresponds to a finite memory window, in the time domain. Hence, we expect that the effective power $P_{eff}(t)$ to be a filtered version of the transmitted power $P(t)$ over such a time window. For instance, in the limit of a very narrow time window, $P_{eff}(t)$ should coincide with the local power at time $t$ (thus $P_{eff}(t_s) = P_{peak}$), while for a very large memory
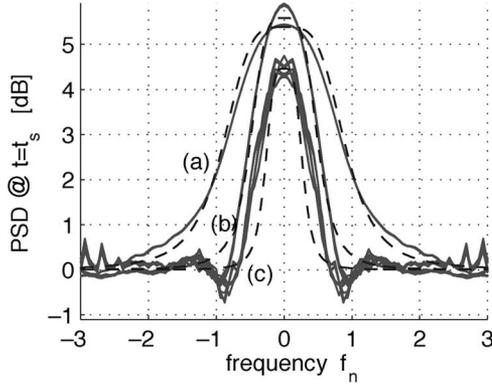
Fig. 6. Normalized ASE PSD versus frequency $f_n$ for a $20 \times 100$ km RZ-DPSK system with $z/L_d = 0$, $\overline{\Phi}_{NL} = 0.3\pi$, and (a) $S = 0.03$, (b) $0.1$, (c) $0.35$. Dashed line: CW-equivalent model; solid lines: exact model [12].

time window, the noise interacts with the average power on such a window (thus, $P_{eff}(t_s) = \overline{P}$). For the terrestrial kernel (11), we empirically found that a proper windowing filter for a fully compensated system is

$$H(\omega) = \frac{1}{1 + \left(\frac{S}{4}\omega^2\right)^2} \quad (14)$$

so that $P_{eff}(t) = P(t) \otimes h(t)$, where $h(t)$ is the inverse Fourier transform of $H(\omega)$, while $\otimes$ denotes convolution. Hence, the quasi-stationary ASE noise PSD at time $t_s$ can still be evaluated from (28), in which we use $P_{eff}(t_s)$ in place of $P_{peak}$ in the evaluation of $\Phi_{NL}$. The periodicity of $P_{eff}(t)$, due to the periodicity of the PSK signal power, agrees with the observed cyclostationarity of the ASE noise shown in Fig. 5. In Fig. 6, we compare the exact ASE PSDs [12] at times $t_s$ with our CW-equivalent ASE PSD for an RZ-DPSK $20 \times 100$ km DM system with duty cycle $d = 0.5$ and $\overline{\Phi}_{NL} = 0.3\pi$. The tested strengths are $S = 0.03, 0.1, 0.35$, which correspond to bit-rates $R \cong 5, 10, 20$ Gb/s when $D_{tx} = 8$ ps/nm/km. Note that the PSD at $f_n = 0$ is now a function of strength $S$, and decreases from small $S$ to large $S$, as predicted by (14).

## V. BER EVALUATION

The assumption of Gaussian statistics for the ASE noise before detection allows us to compute the BER of DPSK/DQPSK systems by means of a known KL algorithm for quadratic detectors in Gaussian noise [9], which we suitably modified for PSK modulation formats, much as in [10], and extended them to include PG. For an accurate BER evaluation the KL algorithm is mandatory, since a Gaussian approximation for the electrical sampled current statistics has been shown to fail for the DPSK format [11].

### A. DPSK Case

*1) No PG:* Let us start with the DPSK modulation in absence of PG. Fig. 1(a) depicts the DPSK receiver, where we imagine to collect all amplifiers' noise into a white Gaussian source, added to the optical signal $s(t)$ (obtained through the noiseless SSFM) before the optical filter with frequency

response $H_o(f)$. The one-sided PSD of each noise polarization $N_0$ can be expressed as a function of OSNR from (13). As in [9], for OOK formats, the signal is expanded in Fourier series on its period $N_{bit}T$ as $s(t) = \Sigma_{l=-\bar{L}}^{\bar{L}} s_l e^{j2\pi l(t/N_{bit}T)}$, where $N_{bit}$ is the number of transmitted bits, and $\bar{L}$ takes into account the nonnegligible harmonics of $s(t)$. Note that, in this section, time and frequency are not normalized. Besides, a KL expansion on the Fourier orthonormal basis is applied to the noise around the sampling time $t_k = t_0 + (k-1)T$, $k = 1, \ldots, N_{bit}$, so that the noise copolarized with the signal is[2] $w(t) = \Sigma_{i=-\infty}^{\infty} w_i e^{j2\pi i(t-t_k+T_0)/T_0}$ with $t_k - T_0 < t < t_k$. $T_0$ is a measurement time longer than the receiver memory time, which in practice can be assumed as finite and equal to

$$T_0 = \mu \left( \frac{1}{B_o} + \frac{1}{B_e} \right) + T$$

where $\mu$ plays the role of time-expansion factor, as defined in [9], and $T$ is the memory time of the Mach–Zehnder delay demodulator. The Fourier coefficients $w_i$ are complex independent identically distributed Gaussian random variables (RVs) with zero mean and real and imaginary components of variance $\sigma_w^2 = N_0/(2T_0)$. After the optical filter, the Fourier series of the noise is $n(t) = \Sigma_{i=-M}^{M} n_i e^{j2\pi i(t-t_k+T_0)/T_0}$, where $M$ accounts for the nonnegligible harmonics of $w(t)$ selected by the optical filter and $n_i = w_i h_i$, with $h_i = H_o(i/T_0)$, $i = -M, \ldots, M$, are still Gaussian RVs. We also define the optical signal at the Mach–Zehnder input as $e(t) = s(t) + n(t)$, yielding a photodetected current before the electrical filter $H_R(f)$ equal to

$$y(t) = \frac{1}{4}|e(t) + e(t-T)|^2 - \frac{1}{4}|e(t) - e(t-T)|^2$$
$$= \text{Re}\{e(t)e^*(t-T)\}. \quad (15)$$

The signal at the sampler is $I(t) = y(t) \otimes h_R(t)$, where $h_R(t)$ is the inverse Fourier transform of $H_R(f)$. The insertion of the Fourier series expansion of $s(t)$ and $n(t)$ in $I(t)$ yields a quadratic form for the sampled current $I(t_k)$ at time $t_k$ [9], which can be expressed in a vector notation as

$$I(t_k) = c_k + \mathbf{n}^\dagger \mathbf{v_k} + \mathbf{v_k}^\dagger \mathbf{n} + \mathbf{n}^\dagger \mathbf{T} \mathbf{n} \quad (16)$$

where $\dagger$ indicates transpose conjugate, $k = 1, \ldots, N_{bit}$, and
1) $c_k$ is the sampled current at time $t_k$ in absence of noise;
2) $\mathbf{n} = (n_{-M}, \ldots, n_M)^T$ is a $(2M+1)$-column vector of the Gaussian RVs $n_i$, where T indicates the transpose operation;
3) $\mathbf{T} = (1/2)(\mathbf{D}^\dagger \mathbf{Q} + \mathbf{Q}^\dagger \mathbf{D})$ is a $(2M+1) \times (2M+1)$ matrix that accounts for the noise-to-noise beat, where the matrix $\mathbf{Q}$ has components[3] $q_{im} = H_R((m-i)/T_0)$, with $i, m = 1, \ldots, 2M+1$, and $\mathbf{D}$ is a diagonal matrix of elements $d_{ii} = e^{-j2\pi(i-M-1)T/T_0}$;

---

[2]In the following calculation, we neglect the impact of the noise orthogonal polarization. We will discuss its inclusion at the end of the analysis.

[3]Note that in [9], the indexes $m$ and $i$ were erroneously inverted. The typo was corrected in E. Forestieri, *J. Lightwave Technol.*, vol. 21, p. 1592, Jun. 2003.

4) $\mathbf{v_k} = (1/2)(\mathbf{D}^\dagger \mathbf{x_k} + \mathbf{x_{k-1}})$ is a $(2M+1)$-column vector which accounts for the received signal at time $t_k$, while $\mathbf{x_k}$ has components $x_{ki} = \Sigma_{l=-\bar{L}}^{\bar{L}} \hat{s}_{li} e^{j2\pi(kl/N_{\text{bit}})}$, $i = 1, \dots, 2M+1$, where

$$\hat{s}_{li} = s_l H_R((l/N_{\text{bit}}T) - ((i-M-1)/T_0)) e^{j2\pi l(t_o/N_{\text{bit}}T)}.$$

The elements $x_{ki}$ and $\hat{s}_{li}$ can be viewed as the components of the $N_{\text{bit}} \times (2M+1)$ matrix $\mathbf{X}$ and $(2\bar{L}+1) \times (2M+1)$ matrix $\widehat{\mathbf{S}}$, respectively. By this way, we have that $\mathbf{X} = \widehat{\mathbf{F}} \cdot \widehat{\mathbf{S}}$, where the $N_{\text{bit}} \times (2\bar{L}+1)$ matrix $\widehat{\mathbf{F}}$ has components $e^{j2\pi kl/N_{\text{bit}}}$. This matrix product can be quickly evaluated by observing that it is $\widehat{\mathbf{F}} = \mathbf{F} \cdot \widehat{\mathbf{I}}$, being $\mathbf{F}$ the $N_{\text{bit}} \times N_{\text{bit}}$ Fourier matrix whose $(k, l)$ element is still equal to $e^{j2\pi lk/N_{\text{bit}}}$, and $\widehat{\mathbf{I}} = [\mathbf{I}, \mathbf{I}, \dots, \mathbf{I}]$ is a repetition of $N_T = (2\bar{L}+1)/N_{\text{bit}}$ instances of the identity matrix.[4] Note that $N_T$ is the number of discrete points per bit of $s(t)$. We finally have $\mathbf{X} = \mathbf{F} \cdot (\widehat{\mathbf{I}} \cdot \widehat{\mathbf{S}})$. The first product can be quickly evaluated by the fast Fourier transform (FFT) algorithm, while the second involves only $N_T$ nonzero multiplications per row, which implies an under sampling of $s(t)$.

The insertion of the optical-filtering transformation $\mathbf{n} = \mathbf{H_o}\mathbf{w}$ in (16), with $\mathbf{H_o} = \text{diag}\{\text{h}_i\}$, yields $\mathbf{n}^\dagger \mathbf{Tn} = \mathbf{w}^\dagger \mathbf{H_o^\dagger T H_o}\mathbf{w} = \mathbf{w}^\dagger \mathbf{A}\mathbf{w}$, where the matrix $\mathbf{A}$ is Hermitian. $\mathbf{A}$ can be diagonalized by the unitary matrix $\mathbf{U}$ formed by its eigenvectors, corresponding to the eigenvalues $\lambda_i$, $i = 1, \dots, 2M+1$. Since a unitary transformation does not change the statistics of the white Gaussian vector $\mathbf{w}$, after diagonalization one can express the sampled current as a sum of independent chi-square RVs. Strictly speaking, the introduction of the whitened vector $\mathbf{z} \triangleq \mathbf{U}^\dagger \mathbf{H_o^{-1}}\mathbf{n}$ allows us to write the received current in (16) as

$$I(t_k) = c_k + \Sigma_{i=1}^{i_M} \lambda_i |z_i + (b_{ki}/\lambda_i)|^2 - \Sigma_{i=1}^{i_M} (|b_{ki}|^2/\lambda_i)$$

where $i_M = 2M+1$ and, for a fixed $k$, $b_{ki}$, $i = 1, \dots, i_M$ are the components of the complex vector $\mathbf{b_k} = \mathbf{U}^\dagger \mathbf{H_o^\dagger}\mathbf{v_k}$. The covariance matrix of $\mathbf{z}$ is $\text{E}\{\mathbf{z} \cdot \mathbf{z}^\dagger\} = 2\sigma_w^2 \mathbf{I}$; hence, $z_i$ are independent identically distributed complex Gaussian RVs. The RV $m_k = \Sigma_{i=1}^{i_M} \lambda_i |z_i + (b_{ki}/\lambda_i)|^2$ is a noncentral quadratic form of Gaussian RVs with a moment generating function (MGF) equal to [25], [26]

$$\Psi_{m_k}(s) = \prod_{i=1}^{i_M} \frac{\exp\left(\frac{\alpha_{ki} s}{1-\beta_i s}\right)}{(1-\beta_i s)^{\xi/2}} \tag{17}$$

where $\alpha_{ki} = |b_{ki}|^2/\lambda_i$ and $\beta_i = 2\lambda_i \sigma_w^2$, and $\xi = 2$ since the $z_i$ are complex RVs. From the MGF of $m_k$, one can finally evaluate the BER through a saddle point integration, as described in detail in [9]. Notice that the BER should not be derived from the inverse FFT of the MGF because the tails of the corresponding pdf are limited by a numerical precision. If there is ASE in the orthogonal polarization with Fourier coefficients

$\hat{\mathbf{n}}$, an additional quadratic form $\hat{\mathbf{n}}^\dagger \mathbf{T}\hat{\mathbf{n}}$ should be added to (16). This new term leads to an MGF as before but by using $\xi = 4$.

*2) With PG:* As shown in the Appendix, the inclusion of noise PG breaks the symmetry between the in-phase and quadrature noise PSD, which calls for doubling the dimension of $\mathbf{w}$. Moreover, (29) reveals that PG acts as a transfer matrix $\mathbf{H}$ on the noise $w(t)$. These conditions ensure the reliability of the above procedure, even with PG, provided that $n(t)$ is whitened.

To this aim, we decompose the vectors $\mathbf{n}$ and $\mathbf{w}$ in their real and imaginary components, e.g., $\mathbf{n} = \mathbf{n_r} + j\mathbf{n_i}$, which we compact in the vector notation into $\mathbf{n_d} = [\mathbf{n_r}; \mathbf{n_i}]$ and $\mathbf{w_d} = [\mathbf{w_r}; \mathbf{w_i}]$, respectively, with doubled dimension $4M+2$. Equation (16) can be written in terms of $\mathbf{n_d}$ as

$$I(t_k) = c_k + \mathbf{n_d^T}\mathbf{v_d} + \mathbf{v_d^T}\mathbf{n_d} + \mathbf{n_d^T}\mathbf{T_d}\mathbf{n_d} \tag{18}$$

where $\mathbf{v_d} = [\mathbf{v_r}; \mathbf{v_i}]$, and $\mathbf{T_d} = \begin{bmatrix} \mathbf{T_r} & -\mathbf{T_i} \\ \mathbf{T_i} & \mathbf{T_r} \end{bmatrix}$, where the indexes $(r, i)$ indicate the real/imaginary component of $\mathbf{v_k}$ and $\mathbf{T}$, respectively. We need to relate $\mathbf{n_r}, \mathbf{n_i}$ to the corresponding complex Fourier coefficients of the in-phase and quadrature components $\mathbf{n_p}, \mathbf{n_q}$ of $n(t)$. It turns out that by introducing the vectors $\mathbf{n_{PG}} = [\mathbf{n_p}; \mathbf{n_q}]$ and $\mathbf{w_{PG}} = [\mathbf{w_p}; \mathbf{w_q}]$, we get

$$\mathbf{n_{PG}} = \mathbf{H_{PG}}\mathbf{w_{PG}} \triangleq \begin{bmatrix} \mathbf{H_{11}} & \mathbf{H_{12}} \\ \mathbf{H_{21}} & \mathbf{H_{22}} \end{bmatrix} \mathbf{w_{PG}} \tag{19}$$

where $\mathbf{H_{im}} = \text{diag}\{\text{h}_{im}(l/\text{T}_0)\}$, $(i, m) = (1, 2)$, and $l = -M, \dots, M$ includes the effect of the link of the postcompensating fiber and of the optical filter on the noise. The relation between $h_{im}$ and the received PSD is written in (29). Under the assumption of equivalent-stationary noise the PSD is evaluated using $P_{\text{eff}}(t_s)$, which turns out to be the same for all bits due to the periodicity of the DPSK signal. We observe that with PG the GVD of the postcompensating fiber affects the statistics of the colored ASE noise, unlike the case without PG, where the circularly symmetric noise remains unchanged. By writing $\mathbf{n} = \mathbf{n_p} + j\mathbf{n_q}$, and using (21), the vector $\mathbf{n_d}$ can be easily related to $\mathbf{n_{PG}}$ by the relationship $\mathbf{n_d} = \mathbf{B}\mathbf{n_{PG}}$, where the $(4M+2) \times (4M+2)$ matrix $\mathbf{B}$ takes the form

$$\mathbf{B} = \frac{1}{2} \begin{bmatrix} \mathbf{I}+\mathbf{D_o} & j(\mathbf{I}-\mathbf{D_o}) \\ -j(\mathbf{I}-\mathbf{D_o}) & \mathbf{I}+\mathbf{D_o} \end{bmatrix}. \tag{20}$$

In (20), $\mathbf{D_o}$ is a $(2M+1) \times (2M+1)$ matrix equal to one on the off-diagonal elements and 0 otherwise. It follows that $\mathbf{n_d} = \mathbf{B}\mathbf{H_{PG}}\mathbf{B}^{-1}\mathbf{w_d} \triangleq \mathbf{H_B}\mathbf{w_d}$, where $\mathbf{H_B}$ is a real matrix. We can repeat the same steps followed in presence of a white noise, obtaining again a quadratic form for $I(t_k)$ but now with $i_M = 4M+2$. The matrix $\mathbf{A}$ is now a symmetric real matrix equal to $\mathbf{H_B^T}\mathbf{T_d}\mathbf{H_B}$, and the whitening operation on the received noise yields $\mathbf{z} \triangleq \mathbf{U}^T \mathbf{H_B^{-1}}\mathbf{n_d}$, whose covariance matrix becomes $\text{E}\{\mathbf{z} \cdot \mathbf{z}^T\} = \sigma_w^2 \mathbf{I}$ here. We finally have the same MGF as in (17) but, now, with $\xi = 1$, since in this context, the $z_i$ components are real RVs. $\beta_i$ and $\alpha_{ki}$ are computed by using the new eigenvalues and eigenvectors of matrix $\mathbf{A}$ with PG.

In this case, the ASE in the orthogonal polarization requires the replacement of $(1 - \beta_i s)$ with $(1 - \beta_i s)(1 - \hat{\beta}_i s)$ in the

---

[4]Since $N_{\text{bit}}$ is usually a power of 2, for $N_T$ to be an integer, one should choose an expansion of $s(t)$ over $2\bar{L}$ frequencies (instead of $2\bar{L}+1$) and choose $\bar{L}$ as a power of 2.
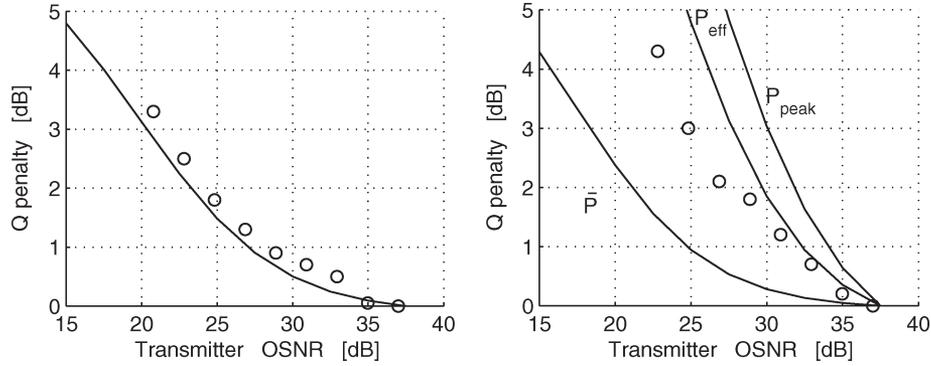
Fig. 7. $Q$-penalty versus transmitter OSNR for NRZ- (left) and RZ-DPSK (right) for the experimental system tested by Kim and Gnauck in [13] with a launched power of 7 dBm and $d = 0.33$.

denominator of the main fraction only, with $\hat{\beta}_i$ evaluated in absence of PG.

### B. DQPSK Case

Fig. 1(b) shows the DQPSK receiver. We infer from it that the insertion of the phase shifts of $\pi/4$ or $-(\pi/4)$, respectively, on the upper and lower arm of the receiver, gives an output signal in absence of noise equal to $\cos(\Delta\Phi_k - (\pi/4))$ and $\sin(\Delta\Phi_k - (\pi/4))$, where $\Delta\Phi_k$ is the differential optical phase at the sampling time $t_k$. Because the decision symbol is equal to $\pm 1/\sqrt{2}$, we identify two DPSK modulated signals on each receiver's arm, $I_P(t_k) \triangleq I_k$ and $I_Q(t_k) \triangleq Q_k$, respectively, which allows us to apply the extended KL method previously described, both with and without PG. Hence, $I_k$ and $Q_k$ can still be written as quadratic forms as in Section V-A, but now, the signal vector is equal to $\mathbf{v_k} = (1/2) \times (\mathbf{D}^\dagger \mathbf{x_k} e^{\mp j(\pi/4)} + \mathbf{x_{k-1}} e^{\pm j(\pi/4)})$, and the noise-to-noise beat matrix becomes $\mathbf{T} = (1/2)(\mathbf{D}^\dagger \mathbf{Q} e^{\mp j(\pi/4)} + \mathbf{Q}^\dagger \mathbf{D} e^{\pm j(\pi/4)})$, where the upper/lower sign before $\pi/4$ in the exponential terms refers to the computation of $I_k/Q_k$. In general, the RVs $I_k$ and $Q_k$ are not independent, but, as an average case, we compute the BER of the DQPSK signal as if they were [27]. The overall BER is defined as

$$\text{BER} = (1/2)[1 - (1 - P_Q)(1 - P_P)] \simeq (1/2)(P_Q + P_P)$$

where $P_Q$ and $P_P$ are the BER on $Q_k$ and $I_k$, respectively. Alternatively, $P_Q$ and $P_P$ can be separately studied [28].

## VI. RESULTS AND DISCUSSION

In this section, we prove the accuracy of our model for BER evaluation in the presence of PG by checking it against experimental and numerical results. We then apply it to the analysis of PSK system performance, comparing binary and quaternary schemes.

We first tried to reproduce the experimental results of Kim and Gnauck in [13] for a $6 \times 100$ km nonzero dispersion-shifted link working at 10 Gb/s and with a launched power of 7 dBm. Fig. 7 shows the $Q$-factor penalty measured in [12] with circles and the prediction of our model (solid line) for (left) NRZ- and (right) RZ-DPSK ($d = 0.33$). Note that for
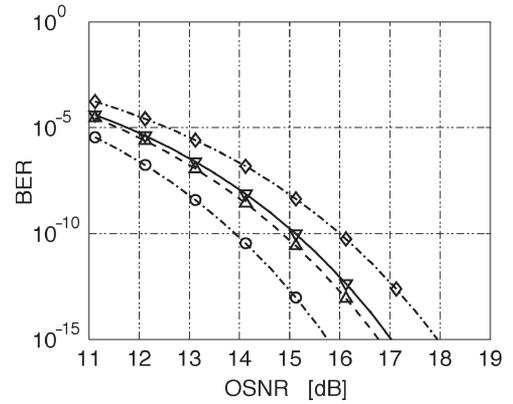


Fig. 8. BER versus OSNR for a 20 Gb/s 20-span full compensated RZ-DPSK system ($d = 0.5$) with $\overline{\Phi}_{\text{NL}} = 0.2\pi$. Triangles up: BER using the algorithm in [12]. Triangles down: proposed model by using $P_{\text{eff}}$. Circles/Diamond: proposed model with $\overline{P}$ or $P_{\text{peak}}$.

our method, the $Q$-factor is evaluated by inverting the BER. In each curve, the penalty is referred to the $Q$-factor at a transmitter OSNR = 37 dB, which is the highest value used in the experiment. For the details of the system setup, see [13]. For the RZ case, we evaluated the $Q$-penalty by using the CW-equivalent ASE model with either the average power $\overline{P}$, or the peak power $P_{\text{peak}}$, or the effective power $P_{\text{eff}}$ obtained through (14). From the comparison with experimental data, the case using $P_{\text{eff}}$ is found to reasonably fit the experimental data up to penalties of 2 dB, with some over estimation at lower values of OSNR.

In Fig. 8, we also numerically tested our BER algorithm using the CW-equivalent ASE with either $\overline{P}$ (circles), or $P_{\text{peak}}$ (diamonds) or $P_{\text{eff}}$ (up-triangles) by comparison with the exact BER (down triangles) evaluated through the algorithm proposed in [12]. The BER curves are plotted against the received OSNR and refer to a 20-span fully compensated RZ-DPSK ($d = 0.5$) system, with a transmission fiber chromatic dispersion $D_{\text{tx}} = 8$ ps/nm/km, pre- and postfibers optimized, for $S = 0.35$ (i.e., $R = 20$ Gb/s) and for an average $\overline{\Phi}_{\text{NL}} = 0.2\pi$. The best fit is again given by the case using $P_{\text{eff}}$, while the use of $\overline{P}$ or $P_{\text{peak}}$ under/overestimates the BER.

We now use our BER algorithm for the numerical study of both NRZ/RZ DPSK and DQPSK systems, according to the fully compensated schemes explained in Section II. The performance is measured in terms of OSNR penalty (at
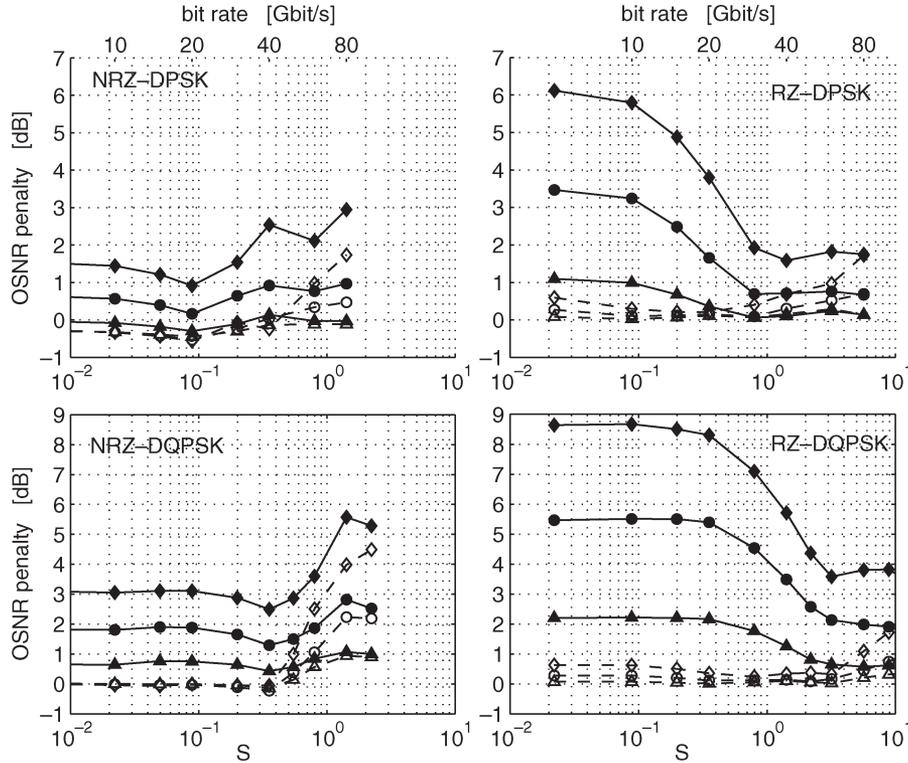
Fig. 9.   OSNR penalty versus $S$ for (left) NRZ and (right) RZ ($d = 0.5$) (top) DPSK and (bottom) DQPSK modulation formats with (solid line) and without PG (dashed line). Triangles: $\overline{\Phi}_{\mathrm{NL}} = 0.1\pi$; Circles: $\overline{\Phi}_{\mathrm{NL}} = 0.2\pi$; Diamonds: $\overline{\Phi}_{\mathrm{NL}} = 0.3\pi$.

BER $= 10^{-10}$) referred to back to back at different nonlinear phases and strengths $S$. For the DQPSK case, $L_\Delta$(and thus $S$) is referred to the bit rate of the signal.

In Fig. 9, we plot the OSNR penalty versus $S$ for $\overline{\Phi}_{\mathrm{NL}} = 0.1\pi$ (triangles), $\overline{\Phi}_{\mathrm{NL}} = 0.2\pi$ (circles), and $\overline{\Phi}_{\mathrm{NL}} = 0.3\pi$ (diamonds) computed with (solid line) and without (dashed line) PG, for NRZ- (left) and RZ-(right) DPSK (top) and DQPSK (bottom) signals. We also draw in the top $x$-axis the bit rate corresponding to $S$ for the system under investigation ($D_{\mathrm{tx}} = 8$ ps/nm/km, $d = 0.5$). The pre- and postcompensation fibers were optimized for each nonlinear phase in the presence of PG, and the same values were used even in absence of PG. We numerically searched the optimal pre- within the normalized range $[-5S, 5S]$ and the optimal post- by varying the overall normalized cumulated dispersion, i.e., the sum of the pre- and postcompensation, within the range $[-0.3, 0.3]$. The optimal normalized precompensation was found to be very close to $-S \log(2)$, much like in OOK systems [29], while the optimal total dispersion was close to zero for all tested nonlinear phase values (hence, postcompensation values were close to $+S \log(2)$), which is strikingly different from OOK systems. In all cases, the performance is set by PG at small $S$, while the distorting nonlinear effect of self-phase modulation (SPM) is dominant at large $S$. For this reason, at each $\overline{\Phi}_{\mathrm{NL}}$, the distance between the curves with and without PG decreases for increasing $S$, both in DPSK and DQPSK. However, RZ pulse shaping is more robust to ISI both with and without PG at large $S$, while at small $S$, it suffers more from PG than NRZ shaping [13]. This last conclusion should be expected since, when PG is dominant, i.e., for small $S$, in the RZ system, $P_{\mathrm{eff}}$ approaches

$P_{\mathrm{peak}}$, which is twice that of NRZ shaping. Regarding the comparison between binary and quaternary systems, we first note that the reduced distance among the DQPSK symbols decreases its robustness to SPM with NRZ pulses, while the use of RZ pulses can actually reduce such an impairment. On the other hand, PG strongly worsens DQPSK performance, much more than in the DPSK case. The reason is that the amount of PG depends solely on the symbol rate, regardless of the bit rate. Hence, since quaternary formats support two channels at a halved symbol rate, they are less robust to PG, where PG is much stronger at lower symbol rates.

It is worth observing that the conclusions drawn from Fig. 9 may change with different dispersion maps, i.e., with a nonzero in-line dispersion. The presence of a practical in-line dispersion should not affect the validity of the Gaussian assumption for the tested nonlinear phases, according to what observed in [16] and [30]. However, the in-line dispersion is expected to shift the effective power $P_{\mathrm{eff}}$ away from the value predicted by (14), since it affects the system memory time. Such an issue is worth of further investigation.

### A. Comparison With OOK

We compared the $Q$-factor of DPSK format with the one of an OOK modulation with NRZ pulses. For OOK, we evaluated the BER by following the same lines shown in Section V-A adapted to an OOK demodulator. A peculiar difference between the models relies on accounting PG for OOK only during marks. Moreover, since PG for OOK appears only at large nonlinear phases [16], we estimated its PSD through off-line
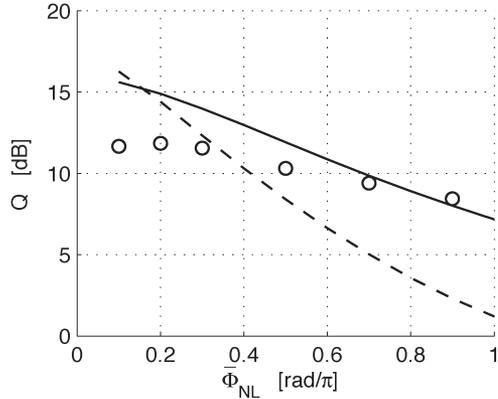
Fig. 10. $Q$-factor versus $\overline{\Phi}_{\mathrm{NL}}$ for NRZ-DPSK (solid), RZ-DPSK at 50% duty cycle (dashed), and NRZ-OOK (circles). $20 \times 100$ km fully compensated system at $R = 10$ Gb/s ($S = 0.02$ for $D_{\mathrm{tx}} = 8$ ps/nm/km), OSNR = 11 dB/0.1 nm.

Monte Carlo simulations as in [16] and [30] to correctly capture the strongly nonlinear PG. For DPSK, instead, we used the analytical ASE PSD (28) since the performance rapidly deteriorates already at small nonlinear phases. For this reason, penalties are slightly overestimated at large $\overline{\Phi}_{\mathrm{NL}}$, where the system is however unable to work.

The optical link analyzed is the same as in Fig. 2, working at an OSNR = 11 dB/0.1 nm. Fig. 10 depicts the measured $Q$-factor versus the average nonlinear phase $\overline{\Phi}_{\mathrm{NL}}$ defined in (1), for NRZ-DPSK (solid), RZ-DPSK (dashed) and NRZ-OOK (circles). From the figure, we note the well-known 3-dB difference between DPSK and OOK at small $\overline{\Phi}_{\mathrm{NL}}$, while for increasing $\overline{\Phi}_{\mathrm{NL}}$, we observe different distortions on the modulation formats, mainly due to PG [30]. As discussed in Section IV-A, RZ-DPSK is strongly affected by PG since in this case $P_{\mathrm{eff}}$ is close to $P_{\mathrm{peak}}$, twice the value than in NRZ-DPSK. Meanwhile, we observe that OOK for increasing $\overline{\Phi}_{\mathrm{NL}}$ recovers the previously mentioned 3-dB gap, and it overcomes DPSK at large $\overline{\Phi}_{\mathrm{NL}}$. The reason can be found in the greater PG-induced increase of the noise quadrature component with respect to the in-phase one. Such an inflation has a strong impact in terms of nonlinear phase noise on the performance of DPSK.

## VII. CONCLUSION

A PG approach to the ASE noise impact on DM PSK modulated systems in the nonlinear regime has been presented. We have shown that in practical DM systems in presence of a chromatic dispersion, the statistics of the received ASE noise can still be assumed as Gaussian, and a novel closed-form expression for the received ASE PSD has been obtained from a linearization of the NLSE based on a CW signal (28). We have shown that, even in presence of the signal modulation, an equivalent stationary ASE process exists, which has the same statistics as the true nonstationary ASE process at the sampling times. The CW-equivalent ASE PSD is obtained by using a lowpass filtered version of the modulating power at the sampling times in the calculation of the nonlinear phase in (28).

Once the Gaussian ASE statistics are available, the BER can be computed by extending a known KL method for quadratic detectors in Gaussian noise. Such an approach avoids the

explicit calculation of the nonlinear phase noise statistics, and allows taking ISI arising from the nonlinear waveform distortion, optical, and electrical filtering into account. Our method has shown good agreement with both published experimental results and exact numerical results obtainable from the method of Holzlöhner *et al.* [12] but with very large savings in computation time.

We finally applied our model to a performance evaluation of single-channel DPSK and DQPSK systems with both NRZ and RZ supporting pulses, for varying map strength and nonlinear phase values, both in the presence of PG and SPM distortion. We found that PG is the main impairment at small map strengths, especially for RZ pulses. We also found that the doubled spectral efficiency of DQPSK is paid in terms of a larger penalty due to PG. In a comparison with a single-channel NRZ-OOK at 10 Gb/s, we found that the 3-dB OSNR advantage of DPSK disappears at large nonlinear phases due to the larger sensitivity of DPSK to nonlinear phase noise.

## APPENDIX
## SOLUTION OF (10)

In Section IV, in the case of a CW signal, we derived the averaged (10), which describes the noise dynamics over a slow spatial scale. To solve such an equation, we now decompose the perturbation field in its real and imaginary components, i.e., $a(z,t) = a_{\mathrm{p}}(z,t) + ja_{\mathrm{q}}(z,t)$. Having referred the noise to the signal nonlinear phase [see (4)], $a_{\mathrm{p}}$ and $a_{\mathrm{q}}$ coincide with the in-phase and quadrature noise components. In the Fourier domain, we have $\tilde{a}(z,\omega) = \tilde{a}_{\mathrm{p}}(z,\omega) + j\tilde{a}_{\mathrm{q}}(z,\omega)$, where

$$\tilde{a}_{\mathrm{p}}(z,\omega) = \frac{\tilde{a}(z,\omega) + \tilde{a}^*(z,-\omega)}{2}$$

$$\tilde{a}_{\mathrm{q}}(z,\omega) = \frac{\tilde{a}(z,\omega) - \tilde{a}^*(z,-\omega)}{2j} \qquad (21)$$

with which we form the column vector $\widetilde{\mathbf{a}} \triangleq [\tilde{a}_{\mathrm{p}}, \tilde{a}_{\mathrm{q}}]^{\mathrm{T}}$. In the same way, we write the Fourier transform of the noise $\widetilde{W}(z,\omega) = \widetilde{W}_{\mathrm{p}}(z,\omega) + j\widetilde{W}_{\mathrm{q}}(z,\omega)$ and form the vector $\widetilde{\mathbf{W}} \triangleq [\widetilde{W}_{\mathrm{p}}, \widetilde{W}_{\mathrm{q}}]^{\mathrm{T}}$. Substituting (21) in (10) yields the following matrix differential equation:

$$\frac{\partial \widetilde{\mathbf{a}}(z,\omega)}{\partial z} = \mathbf{M}(\omega) \cdot \widetilde{\mathbf{a}}(z,\omega) + \widetilde{\mathbf{W}}(z,\omega) \qquad (22)$$

with initial condition $\widetilde{\mathbf{a}}(0,\omega) = \mathbf{0}$ and with

$$\mathbf{M}(\omega) = \frac{\omega^2}{2L_{\mathrm{d}}} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \mathcal{R}(0) \begin{bmatrix} r_i & 1 - r_r \\ -1 - r_r & -r_i \end{bmatrix} \qquad (23)$$

where $r_r$ and $r_i$ are the real and imaginary parts of the kernel $r(\omega)$, respectively. Thanks to the method of averaging, (23) is a linear differential equation with constant coefficients, whose solution is [18]

$$\tilde{\mathbf{a}}(z,\omega) = \int_0^z e^{\mathbf{M}(\omega)(z-x)} \widetilde{\mathbf{W}}(x,\omega) dx \qquad (24)$$

where the matrix exponential can be decomposed as

$$e^{\mathbf{M}(\omega)z} = \cosh(kz)\mathbf{I} + \frac{\sinh(kz)}{kz}\mathbf{M}z$$

where $\mathbf{I}$ is the identity matrix and $\pm k$ the eigenvalues of matrix $\mathbf{M}$, with

$$kz = \sqrt{\Phi_{\mathrm{NL}}^2(z)\,|r(\omega)|^2 - \left(\frac{\omega^2}{2}\frac{z}{L_{\mathrm{d}}} + \Phi_{\mathrm{NL}}(z)\right)^2}$$

where, at the space coordinate $z$, $\Phi_{\mathrm{NL}}(z) = z\mathcal{R}(0)$ accounts for the cumulated nonlinear phase by the CW.

By applying the Wiener–Kinchine theorem [15], we can now evaluate the PSD matrix of the received ASE noise in (24) as

$$\mathbf{G}(z,\omega) = \begin{bmatrix} G_{\mathrm{pp}} & G_{\mathrm{pq}} \\ G_{\mathrm{qp}} & G_{\mathrm{qq}} \end{bmatrix} \triangleq \lim_{\tau \to \infty} \frac{1}{\tau} E\left\{\tilde{\mathbf{a}}(z,\omega)\tilde{\mathbf{a}}^\dagger(z,\omega)\right\} \quad (25)$$

where $\tau$ is a time window over which we truncate the ASE terms $W_{\mathrm{p}}(z,t)$ and $W_{\mathrm{q}}(z,t)$. $G_{\mathrm{pp}}$ and $G_{\mathrm{qq}}$ are, respectively, the PSD of the in-phase and quadrature components of $\tilde{\mathbf{a}}$, while $G_{\mathrm{pq}} = G_{\mathrm{qp}}^*$ is the cross spectrum. From (24) and noting that $(e^{\mathbf{M}z})^\dagger = e^{\mathbf{M}^\dagger z}$, we get

$$\mathbf{G} = \int_0^z\int_0^z e^{\mathbf{M}(z-y)}E\left\{\widetilde{\mathbf{W}}(y,\omega)\widetilde{\mathbf{W}}^\dagger(x,\omega)\right\}e^{\mathbf{M}^\dagger(z-x)}dy\,dx. \quad (26)$$

Because $\widetilde{\mathbf{W}}$ is the white noise with PSD per component per unit length equal to $\sigma^2$, we have $E\{\widetilde{\mathbf{W}}(y,\omega)\widetilde{\mathbf{W}}^\dagger(x,\omega)\} = \sigma^2\delta(y-x)\mathbf{I}$, which, inserted into (26), gives

$$\mathbf{G} = \sigma^2\int_0^z e^{\mathbf{M}\eta}e^{\mathbf{M}^\dagger\eta}d\eta. \quad (27)$$

Substituting (23) in (27) yields the following ASE PSD normalized to $\sigma^2 z$, i.e., the white noise PSD per component at coordinate $z$:

$$\frac{\mathbf{G}}{\sigma^2 z} = k_0\mathbf{I} - k_1\begin{bmatrix} -r_i & r_r \\ r_r & r_i \end{bmatrix} - k_2\begin{bmatrix} r_r & r_i \\ r_i & -r_r \end{bmatrix} \quad (28)$$

where

$$\begin{cases} k_0 = 1 + 4\Phi_{\mathrm{NL}}^2(z)\frac{|r|^2}{(2kz)^2}\left(\frac{\sinh 2kz}{2kz} - 1\right) \\ k_1 = 2\Phi_{\mathrm{NL}}(z)\frac{\cosh 2kz - 1}{(2kz)^2} \\ k_2 = 4\Phi_{\mathrm{NL}}(z)\left(\frac{z}{L_{\mathrm{d}}}\frac{\omega^2}{2} + \Phi_{\mathrm{NL}}(z)\right)\frac{1}{(2kz)^2}\left(\frac{\sinh 2kz}{2kz} - 1\right). \end{cases}$$

If the link is followed by a linear device, e.g., an ideal compensating fiber and/or an optical filter, the output PSD can be found by a straightforward extension of (25) [15, p. 319]. In the limit $\omega \longrightarrow 0$, we have $(G_{\mathrm{pp}}/\sigma^2 z) \longrightarrow 1$ and $(G_{\mathrm{qq}}/\sigma^2 z) \longrightarrow 1 + (4/3)\Phi_{\mathrm{NL}}^2$; hence, at large $\Phi_{\mathrm{NL}}$, the quadrature component is dominant and manifests itself as nonlinear phase noise.

Since $\mathbf{G}$ is a positive semidefinite matrix, it has the representation $\mathbf{G} = \sigma^2 z(\mathbf{H} \cdot \mathbf{H}^\dagger)$ [25], where

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} \sqrt{G_{\mathrm{pp}}} & 0 \\ \left(\frac{G_{\mathrm{pq}}^*}{\sqrt{G_{\mathrm{pp}}}}\right) & \sqrt{\left|G_{\mathrm{qq}} - \frac{|G_{\mathrm{pq}}|^2}{G_{\mathrm{pp}}}\right|} \end{bmatrix}. \quad (29)$$

This condition ensures that the received noise has the same statistics as a dummy Gaussian white noise with PSD $\sigma^2 z\mathbf{I}$ filtered by the transfer matrix $\mathbf{H}$.

## REFERENCES

[1] A. Gnauck, "Phase shift keyed transmission," in *Proc. OFC*, Atlanta, GA, Mar. 2003, pp. 14–15, Paper ThE1.
[2] O. Vassilieva, T. Hoshida, S. Choudhary, and H. Kuwahara, "Nonlinear tolerant and spectrally efficient 86 Gb/s RZ-DQPSK format for a system upgrade," in *Proc. OFC*, Atlanta, GA, Mar. 2003, pp. 22–24, Paper ThE7.
[3] A. H. Gnauck and P. J. Winzer, "Optical phase-shift-keyed transmission," *J. Lightw. Technol.*, vol. 23, no. 1, pp. 115–130, Jan. 2005.
[4] G. Charlet, E. Corbel, J. Lazaro, A. Klekamp, R. Dischler, P. Tran, W. Idler, H. Mardoyan, A. Konczykowska, F. Jorge, and S. Bigo, "WDM transmission at 6-Tbit/s capacity over transatlantic distance, using 42.7-Gb/s differential phase-shift keying without pulse carver," *J. Lightw. Technol.*, vol. 23, no. 1, pp. 104–107, Jan. 2005.
[5] J.-P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, no. 23, pp. 1351–1353, Dec. 1990.
[6] K.-P. Ho, "Probability density of nonlinear phase noise," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 20, no. 9, pp. 1875–1879, Sep. 2003.
[7] ——. (2005, Sep.). *Statistical Properties of Nonlinear Phase Noise*, vol. 3. Hauppague, NY: Nova Science [Online]. Available: http://arxiv.org/abs/physics/0303090
[8] Y. Yadin, M. Shtaif, and M. Orenstein, "Nonlinear phase noise in phase modulated WDM fiber-optic communications," *IEEE Photon. Technol. Lett.*, vol. 16, no. 5, pp. 1307–1309, May 2004.
[9] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and post-detection filtering," *J. Lightw. Technol.*, vol. 18, no. 11, pp. 1493–1503, Nov. 2000.
[10] J. Wang and J. M. Kahn, "Impact of chromatic and polarization-mode dispersions on DPSK systems using interferometric demodulation and direct detection," *J. Lightw. Technol.*, vol. 22, no. 2, pp. 362–371, Feb. 2004.
[11] G. Bosco and P. Poggiolini, "The impact of receiver imperfections on the performance of optical direct-detection DPSK," *J. Lightw. Technol.*, vol. 23, no. 2, pp. 842–848, Feb. 2005.
[12] R. Holzlöhner, V. S. Grigoryan, C. R. Menyuk, and W. L. Kath, "Accurate calculation of eye diagrams and bit error rates in optical transmission systems using linearization," *J. Lightw. Technol.*, vol. 20, no. 3, pp. 389–400, Mar. 2002.
[13] H. Kim and A. H. Gnauck, "Experimental investigation of the performance limitation of DPSK systems due to nonlinear phase noise," *IEEE Photon. Technol. Lett.*, vol. 15, no. 2, pp. 320–322, Feb. 2003.
[14] A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "New analytical results on fiber parametric gain and its effects on ASE noise," *IEEE Photon. Technol. Lett.*, vol. 9, no. 4, pp. 535–537, Apr. 1997.
[15] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
[16] P. Serena, A. Bononi, J. C. Antona, and S. Bigo, "Parametric gain in the strongly nonlinear regime and its impact on 10 Gb/s NRZ systems with forward-error correction," *J. Lightw. Technol.*, vol. 23, no. 8, pp. 2352–2363, Aug. 2005.
[17] G. P. Agrawal, *Nonlinear Fiber Optics*, 2nd ed. San Diego, CA: Academic, 1995.

[18] D. Zwillinger, *Handbook of Differential Equations*, 3rd ed. Orlando, FL: Academic, 1997.

[19] M. J. Ablowitz and T. Hirooka, "Managing nonlinearity in strongly dispersion-managed optical pulse transmission," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 19, no. 3, pp. 425–439, Mar. 2002.

[20] H. A. Haus, "Quantum noise in a solitonlike repeater system," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 8, no. 5, pp. 1122–1126, May 1991.

[21] A. Cauvin, Y. Frignac, and S. Bigo, "Single-channel nonlinear impairments at various bit-rate in dispersion-managed systems," presented at the Eur. Conf. Optical Communication (ECOC), Rimini, Italy, Sep. 2003, Paper Mo 4.2.4.

[22] P. Serena, A. Orlandini, and A. Bononi, "A parametric gain approach to performance evaluation of DSPK/DQPSK systems with nonlinear phase noise," in *Optical Communication Theory and Techniques*, E. Forestieri, Ed. New York: Springer-Verlag, 2004, pp. 129–136.

[23] M. Midrio, P. Franco, and M. Romagnoli, "Noise statistics in transmission systems with gating dispersion compensation," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 15, no. 11, pp. 2748–2756, Nov. 1998.

[24] M. J. Ablowitz, T. Hirooka, and T. Inoue, "Higher-order analysis of dispersion-managed transmission systems: Solutions and their characteristics," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 19, no. 12, pp. 2876–2885, Dec. 2002.

[25] A. M. Mathai and S. B. Provost, *Quadratic Forms in Random Variables*. New York: Marcel Dekker, 1992.

[26] J. G. Proakis, *Digital Communications*, 4th ed. Singapore: McGraw-Hill, 2001.

[27] S. Benedetto and E. Biglieri, *Principles of Digital Transmission*. New York: Kluwer, 1999.

[28] P. Cho, V. S. Grigoryan, Y. A. Godin, A. Salamon, and Y. Achiam, "Transmission of 25-Gb/s RZ-DQPSK signals with 25-GHz channel spacing over 1000 km of SMF-28 fiber," *IEEE Photon. Technol. Lett.*, vol. 15, no. 3, pp. 473–475, Mar. 2003.

[29] R. I. Killey, H. J. Thiele, V. Mikhailov, and P. Bayvel, "Reduction of intrachannel nonlinear distortion in 40-Gb/s-based WDM transmission over standard fiber," *IEEE Photon. Technol. Lett.*, vol. 12, no. 12, pp. 1624–1626, Dec. 2000.

[30] A. Bononi, P. Serena, J.-C. Antona, and S. Bigo, "Implications of nonlinear interactions of signal and noise in low-OSNR transmission systems with FEC," presented at the Optical Fiber Communication (OFC), Anaheim, CA, Mar. 2005, Paper OTHW5.

**Alessandra Orlandini** was born in Parma, Italy, in December 1973. She received the "Laurea" degree in electronics engineering, major in telecommunications, and the Ph.D. degree in information technology from the Universitá degli Studi di Parma, Parma, Italy, in 1998 and 2002, respectively.

In 2002, she joined the Optical Communication group at the Department of Information Engineering, Universitá degli Studi di Parma, as a Contract Researcher. She is currently collaborating in research projects with Alcatel Corporate Research and Innovations, Marcoussis, France, on long-haul WDM systems. Her research interests are about polarization mode dispersion, its interaction with nonlinearities in WDM systems, and optical differential phase modulation formats in dispersion-managed WDM systems.

**Alberto Bononi** received the Laurea degree (*cum laude*) in electronics engineering from the University of Pisa, Pisa, Italy, in 1988 and the M.A. and Ph.D degrees in electrical engineering from Princeton University, Princeton, NJ, in 1992 and 1994, respectively.

Currently, he is an Associate Professor of Telecommunications at the School of Engineering, Universitá degli Studi di Parma, Parma, Italy. He teaches courses in Probability Theory and Stochastic Processes, Telecommunications Networks, and Optical Communications. In 1990, he worked at General Electric Company (GEC)-Marconi Hirst Research Centre, Wembley, U.K., on a Marconi S.p.A. project on coherent optical systems. From 1994 to 1996, he was an Assistant Professor in the Electrical and Computer Engineering Department, State University of New York (SUNY), Buffalo, teaching courses in Electric Circuits and Optical Networks. In the summers of 1997 and 1999, he was a Visiting Faculty at the Département de Genie Électrique, Université Laval, Laval, QC, Canada, doing research on fiber amplifiers. His present research interests include system design and performance analysis of high-speed all-optical networks, nonlinear fiber transmission for WDM systems, linear and nonlinear polarization mode dispersion, and transient gain dynamics in optical amplifiers.

**Paolo Serena** was born in Piacenza, Italy, in 1973. He received the "Laurea" degree in electronic engineering, major in telecommunications, and the Ph.D. degree in information technology from the Universitá degli Studi di Parma, Parma, Italy, in 1999 and 2003, respectively, for a work on the nonlinear interaction between noise and signal in optical systems.

Currently, he is an Assistant Professor of telecommunications at the Department of Information Engineering, Universitá degli Studi di Parma, where he teaches a course on advanced optical communication. He is currently collaborating in research projects with Alcatel Corporate Research and Innovations, Marcoussis, France, on long-haul wavelength division multiplexing (WDM) systems. His main research interests include the parametric gain (PG) of the amplified spontaneous emission (ASE) noise by nonlinear effects, phase-shifted modulation formats, and performance optimization in the nonlinear regime.