Analysis of Inter-core Crosstalk in Weakly-coupled Multi-core Fiber Coherent Systems

Bruno R. P. Pinheiro, João L. Rebola and Adolfo V. T. Cartaxo

***Abstract*—The influence of the inter-core crosstalk (ICXT) on the performance of weakly-coupled multi-core fiber (WC-MCF) systems with coherent detection and arbitrary inter-core skew is studied numerically and analytically. We analyze the evolution of the instantaneous ICXT power, induced by polarization division multiplexed quadrature amplitude modulation signals, short term average crosstalk (STAXT), and detected ICXT components along the time to get insight on how the ICXT affects the coherent system performance. Numerical results show that, with low skew-symbol rate product ($\ll$1), the STAXT can have large fluctuations along the time and the variance of the detected ICXT can also have large fluctuations along the time, causing large variations of the bit error rate (BER) along time. With large skew-symbol rate product ($\gg$1), the STAXT is almost constant along the time and, although the detected ICXT varies along time, the detected ICXT variance is almost constant along time leading to very small fluctuations of BER along time. Analytical expressions for the variance of the coherently detected ICXT, average BER and optical signal-to-noise ratio (OSNR) penalty are proposed and shown to agree with Monte-Carlo simulation results for arbitrary skew, modulation format order and roll-off factor of the transmitted signals. Numerical results show that the maximum allowable ICXT level for 1 dB OSNR penalty increases when the skew augments and can reach 1.3 dB for a roll-off factor of 1. For most cases of interest of low roll-off factor, the increase of the maximum allowable ICXT level is very small. It is shown that the OSNR penalty estimates agree quite well with other authors' experimental results (with nearly zero roll-off factor).**

***Index Terms*—Bit error rate, coherent detection, inter-core crosstalk, inter-core skew, multi-core fiber, optical signal-to-noise ratio.**

I. Introduction

SINGLE-MODE weakly-coupled multi-core fibers (WC-MCFs) have been proposed to increase the capacity of future optical networks using space division multiplexing [1]–[3]. In order to achieve ultra-high capacity, WC-MCF transmission systems with coherent detection and polarization division multiplexing (PDM) have been reported for long-haul [2], [4], [5], metropolitan and inter-datacenter networks [6].

WC-MCFs can be impaired by the optical coupling between cores, known as inter-core crosstalk (ICXT). The ICXT varies randomly along the fiber [7]–[10], over time [11]–[13] and along frequency [11], [14], [15]. Analytical models have been proposed to describe the stochastic behavior of the ICXT [7]–[15]. Models reported in [9]–[15] generalize the discrete changes model (DCM) reported in [8] to include the stochastic dependence of ICXT along time and frequency, and rely on the fact that the ICXT results mostly from contributions at the phase-matching points (PMPs), i. e. points distributed longitudinally along the fiber, at which the difference between the effective refractive indexes of the interfering and interfered cores is null. The DCMs are particularly interesting for ICXT simulation analyses because they reduce complexity and simulation time to an acceptable level, and take into account specific fiber properties (such as core refractive index profile and radius, pitch) and longitudinal bending and twisting perturbations of the WC-MCF. In particular, the model reported in [12] provides estimates of the autocorrelation function of ICXT power and correlation time that agree with experimental data, meaning that the stochastic time dependence describes the ICXT fluctuations adequately. Additionally, the same model provides estimates of the decorrelation bandwidth in agreement with experimental results, as reported in [15], giving some confidence on its accuracy in describing the dependence of ICXT on the frequency.

The time dependence of the ICXT fluctuations has been studied by analyzing the short-term average crosstalk (STAXT), which is the average ICXT power measured over a short period of time [11]–[13]. Large STAXT fluctuations along time were observed when the ICXT is induced by a continuous wave laser light [11]–[13]. Further analysis of STAXT induced by quadrature amplitude modulation (QAM) signals, typically used in coherent detection systems, showed that, in WC-MCF systems with large inter-core skew or using high symbol rates, the variance of STAXT is very low [13]. This means that, for expected conditions of WC-MCF coherent systems, the STAXT is nearly constant along time and almost constant performance degradation due to ICXT occurs over time [13]. However, as the ICXT instantaneous power fluctuations can achieve a high dynamic range, the relation between the low STAXT variation, the large ICXT instantaneous power fluctuations and their influence on the performance degradation of WC-MCF coherent systems is not thoroughly understood.

Analyses of the impact of ICXT on the performance of WC-MCF coherent systems have already been presented [3], [16]–[20]. A significant increase of the crosstalk-induced performance degradation when the order of the modulation format augments was reported in [16]–[18]. In a recent work [20], the influence of the transmission distance on the ICXT-induced performance degradation in long-haul MCF transmission with coherent detection was studied. In all these works [3], [16]–[20], the influence of the inter-core skew on the ICXT-induced

Manuscript received May XX, XXXX; revised XXXX XX, XXXX. This work was supported in part by Fundação para a Ciência e a Tecnologia (FCT) from Portugal under the project UIDB/50008/2020 and the ISCTE - IUL Merit Scholarship BM-ISCTE-2016.

B. R. P. Pinheiro, J. L. Rebola, and A. V. T. Cartaxo are with the Department of Information Science and Technology, ISCTE - Instituto Universitário de Lisboa, 1649-026 Lisboa, Portugal, and the Instituto de Telecomunicações, 1049-001 Lisboa, Portugal (e-mail: bruno_pinheiro@iscte-iul.pt).

performance degradation was not assessed. Preliminary simulation results reported that the pulse shaping and the inter-core skew influence the variance of the detected ICXT [21]. Thus, a deeper understanding and a rigorous assessment of the impact of the inter-core skew on the performance degradation of WC-MCF coherent systems induced by ICXT is required.

In this work, using the ICXT model proposed in [12], we study numerically and analytically the influence of the ICXT on the performance of the coherently detected PDM-QAM WC-MCF system, taking the inter-core skew into account. We analyze the evolution of STAXT, induced by PDM-QAM signals, the detected ICXT along time, and their relation to the evolution of bit error rate (BER) along time. The main contributions of this work are: (i) a thorough physical insight on the relation between STAXT and detected ICXT, and their influence on the BER are provided for arbitrary inter-core skews; (ii) extension of the study reported in [21] by proposing analytical expressions for the detected ICXT variance and optical signal-to-noise ratio (OSNR) penalty induced by ICXT in WC-MCF coherent systems with arbitrary skew between cores, and comparing with Monte-Carlo (MC) simulation results; (iii) from the derived expressions, the dependence of BER and OSNR penalty on the inter-core skew, pulse shape roll-off factor and QAM modulation format is unveiled and shown to be in agreement with MC simulation results.

The paper is structured as follows. Section II presents the model of the PDM-QAM WC-MCF system with coherent detection. Sections III and IV present the analytical expressions for the coherently detected ICXT variance, average BER and OSNR penalty due to ICXT. In Section V, the relation between STAXT, detected ICXT and BER fluctuations along time is analyzed, and estimates from analytical expressions of the average BER and OSNR penalty are compared with MC simulation results and discussed. Main conclusions are presented in Section VI.

II. WC-MCF Coherent System Model

In order to focus on the effect of the coherent receiver on the ICXT, in this section we consider a single interfering core. The impact of considering multiple interfering cores is discussed in section IV. We assume perfect received signal conditions: (i) the detected signal at the decision circuit input has a raised-cosine (RC) pulse shape to eliminate the inter-symbol interference [22], [23]; (ii) the RC signal is obtained by using root raised-cosine (RRC) filters at the transmitter and at the receiver side to allow a matched filter coherent receiver [22], [23]. Also, linear transmission with propagation delay, group velocity dispersion (GVD) and random polarization rotation is considered in each core. A perfect receiver that compensates for the transmission effects is considered as well. In the following subsections, the models of the dual polarization (DP) transmitter, WC-MCF and coherent receiver are presented.

A. Dual Polarization Transmitter

Perfect DP transmitters that generate PDM-QAM signals with RRC pulse shapes at their outputs are considered. Designating the tested core by n, the interfering core by m, and the DP transmitter that feeds core c by DP-TX$_c$, with $c \in \{m, n\}$, the PDM signal at the output of DP-TX$_c$ (input of core c) can be written as

$$\mathbf{E}_c(t) = E_{c,x}(t)\hat{\mathbf{x}} + E_{c,y}(t)\hat{\mathbf{y}} \tag{1}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ denote the orthogonal polarization direction unity vectors, and $E_{c,p}(t)$, with $p \in \{x, y\}$, is the complex envelope of electrical field of the polarization direction p in core c. The complex envelope of the electrical field at the output of DP-TX$_n$ and polarization $p \in \{x, y\}$ is given by

$$E_{n,p}(t) = \sqrt{P_{n,p}} \sum_{i=-\infty}^{+\infty} a_{p,i}^{(n)} h_{RRC}(t - iT_s) \tag{2a}$$

and at the output of DP-TX$_m$ and polarization $p \in \{x, y\}$ by

$$E_{m,p}(t) = \sqrt{P_{m,p}} \sum_{i=-\infty}^{+\infty} a_{p,i}^{(m)} h_{RRC}(t - iT_s - \tau_{mn}) \tag{2b}$$

where t is the time, T_s is the symbol period, τ_{mn} is the relative time delay between signals at the inputs of cores m and n that models the time misalignment of the signals when they are launched at the fiber cores, $h_{RRC}(t)$ is the RRC pulse shape normalized so that its Fourier transform, $H_{RRC}(f)$, at zero frequency is $H_{RRC}(0) = 1$, and $a_{p,i}^{(c)}$ is the complex amplitude of the i-th transmitted symbol in core c and polarization p. This complex amplitude can be written as $a_{p,i}^{(c)} = a_{I,p,i}^{(c)} + j a_{Q,p,i}^{(c)}$ where $a_{I,p,i}^{(c)}$ and $a_{Q,p,i}^{(c)}$ are, respectively, the amplitudes of the in-phase (I) and quadrature (Q) components of the i-th transmitted symbol in core c and polarization p. For a QAM format with modulation order M (number of distinct symbols), in which $M = 2^{2k}$ with k integer, and with a square constellation, the amplitudes $a_{u,p,i}^{(c)}$, with $u \in \{I, Q\}$, are equally likely zero mean random variables which are uncorrelated for different symbols and for different polarization and quadrature components. The amplitudes of $a_{u,p,i}^{(c)}$ are defined so that their quadratic mean value is $T_s^2/2$. With that normalization of the RRC pulse shape and the amplitudes of the symbols, $P_{n,p}$ and $P_{m,p}$ in eq. (2) are the average powers of polarization p at the output of DP-TX$_n$ and DP-TX$_m$, respectively, with the total transmitted power in core c given by $P_c = P_{c,x} + P_{c,y}$. In our analysis, it is considered that $P_{c,x} = P_{c,y} = P_c/2$, with $c \in \{m, n\}$.

B. WC-MCF model

The ICXT impact on the performance of WC-MCF coherent system is assessed in short time fractions separated by time intervals longer than the ICXT decorrelation time, which is typically in the order of a few minutes [12]. Within each time fraction (with duration much shorter than the decorrelation time), several thousands of QAM symbols are generated to obtain a proper characterization of the STAXT and BER statistics. The number of symbols used depends on the parameter to be characterized and is detailed in Section V. From time fraction to time fraction, the ICXT is uncorrelated, and is evaluated from the dual polarization model of ICXT proposed in [12]. For linear propagation along the WC-MCF, the complex envelope of the electrical field of ICXT induced

in core n along a time fraction by the signal transmitted in the interfering core m can be written as [12]

$$\mathbf{E}_{ICXT}(t) = E_{ICXT,x}(t)\hat{\mathbf{x}} + E_{ICXT,y}(t)\hat{\mathbf{y}} \tag{3}$$

with the field components given by

$$\begin{aligned} E_{ICXT,x}(t) = \mathfrak{F}^{-1}\left[E_{m,x}(\omega)F_{x,x}(\omega)\right] \\ +\mathfrak{F}^{-1}\left[E_{m,y}(\omega)F_{y,x}(\omega)\right] \end{aligned} \tag{4}$$

$$\begin{aligned} E_{ICXT,y}(t) = \mathfrak{F}^{-1}\left[E_{m,x}(\omega)F_{x,y}(\omega)\right] \\ +\mathfrak{F}^{-1}\left[E_{m,y}(\omega)F_{y,y}(\omega)\right] \end{aligned} \tag{5}$$

where ω is the angular frequency, $\mathfrak{F}^{-1}\left[X(\omega)\right]$ is the inverse Fourier transform of $X(\omega)$, and the transfer functions $F_{p,b}(\omega)$, with $b \in \{x, y\}$, are defined as [10], [12]

$$F_{p,b}(\omega) = -j\frac{\overline{K}_{nm}}{\sqrt{2}}e^{-j\overline{\beta}_n(\omega)L}\sum_{k=1}^{N} e^{-j\Delta\overline{\beta}_{mn}(\omega)z_k}e^{-j\phi_{p,b}^{(k)}} \tag{6}$$

where $\overline{K}_{nm}$ is the discrete coupling coefficient, which considers the average inter-core coupling coefficient of both polarization directions [10], and is related to the physical parameters of the MCF [9], $\overline{\beta}_n(\omega)$ is the average of the intrinsic propagation constants of the two polarization directions of core n and takes the propagation delay and GVD effects into account, L is the WC-MCF length, and $\Delta\overline{\beta}_{mn}(\omega)$ is given by [14]

$$\Delta\overline{\beta}_{mn}(\omega) = \Delta\overline{\beta}_{0,mn} + d_{mn}\omega - \Delta D_{mn}\lambda^2\omega^2/(4\pi c) \tag{7}$$

where $\Delta\overline{\beta}_{0,mn}$ is the difference between the averages of the propagation constants at $\omega = 0$, d_{mn} is the average walkoff parameter between core m and n, ΔD_{mn} is the difference between the dispersion parameters of cores m and n, c is the speed of light in a vacuum and λ is the carrier wavelength. The skew between cores m and n, S_{mn}, is given by $d_{mn}L$ [14]. In (6), $\phi_{p,b}^{(k)}$ are random phase shifts (RPSs) that model the random fluctuations in bending radius, twist rate and other physical conditions of the WC-MCF. The RPSs are random variables with uniform distribution between 0 and 2π, and the k-th RPS is introduced at the k-th random coordinate z_k (uniformly distributed between $(k-1)L/N$ and kL/N), with N denoting the number of PMPs. Different RPSs are uncorrelated [11], [12], [14]. The RPSs are uncorrelated from time fraction to time fraction and, as each time fraction duration is much shorter than the ICXT decorrelation time, the RPSs can be considered constant along each time fraction.

As the difference of dispersion parameters is usually small [14], the influence of the difference of dispersion parameters on the amplitude of ICXT field is expected to be reduced in comparison with the skew and GVD influence. Hence, $\Delta D_{mn} = 0$ is considered. Substituting eq. (2b) into eqs. (4) and (5), the electrical field of ICXT at the output of core n in each polarization direction can be written as

$$\begin{aligned} E_{ICXT,p}(t) = -j\frac{\left|\overline{K}_{nm}\right|\sqrt{P_m}}{2}\cdot \\ \sum_{k=1}^{N}\sum_{i=-\infty}^{+\infty}\left[a_{x,i}^{(m)}e^{-j\Phi_{x,p}^{(k)}} + a_{y,i}^{(m)}e^{-j\Phi_{y,p}^{(k)}}\right] \\ \cdot h_{RRC}(t - iT_s - \tau_{mn} - d_{mn}z_k) * \mathfrak{F}^{-1}\{e^{-j\overline{\beta}_n(\omega)L}\} \end{aligned} \tag{8}$$

where $'*'$ stands for convolution and

$$\Phi_{p,b}^{(k)} = \phi_{p,b}^{(k)} + \Delta\overline{\beta}_{0,mn}z_k - \arg\left(\overline{K}_{nm}\right) \tag{9}$$

with $\arg(z)$ the phase of the complex variable z. We stress that, from the viewpoint of the stochastic characterization of ICXT, $\Phi_{p,b}^{(k)}$ has the same stochastic properties as $\phi_{p,b}^{(k)}$. The instantaneous ICXT power, $p_{ICXT}(t)$, at the output of core n is given by

$$p_{ICXT}(t) = \sum_{p\in\{x,y\}}\sum_{u\in\{I,Q\}} E_{ICXT,p,u}^2(t). \tag{10}$$

where $E_{ICXT,p,u}(t)$ are the in-phase ($u = I$) and quadrature ($u = Q$) components of (8) in polarization p.

The STAXT is the average ICXT power measured during a short period of time of duration T (typically 100 to 200 ms [11], [13]), and can be written as [12]

$$STAXT_n(t) = \frac{1}{T}\int_{t-T}^{t} p_{ICXT}(\tau)\, d\tau\,. \tag{11}$$

We stress that the STAXT is different from the mean ICXT power: while the STAXT is a short-term average of ICXT power, the mean ICXT power is a statistical measure.

Considering propagation delay, GVD and random polarization rotation in core n, the electrical field of the signal at the output of core n is, for polarization x, given by [24]

$$\begin{aligned} E'_{n,x}(t) = \left[e^{j\theta}\cos(\Gamma)\cdot E_{n,x}(t)\right. \\ \left.-e^{-j\psi}\sin(\Gamma)\cdot E_{n,y}(t)\right] * \mathfrak{F}^{-1}\left[e^{-j\overline{\beta}_n(\omega)L}\right] \end{aligned} \tag{12}$$

and, for polarization y, by

$$\begin{aligned} E'_{n,y}(t) = \left[e^{j\psi}\sin(\Gamma)\cdot E_{n,x}(t)\right. \\ \left.+e^{-j\theta}\cos(\Gamma)\cdot E_{n,y}(t)\right] * \mathfrak{F}^{-1}\left[e^{-j\overline{\beta}_n(\omega)L}\right] \end{aligned} \tag{13}$$

where the coefficients of electrical field components are the elements of a 2×2 unitary time-varying stochastic Jones matrix $\mathbf{K}$ that describes the random polarization rotation of the electrical field along core n stemming from the birefringence of the MCF. The phases ψ, Γ and θ are random processes that can be assumed constant along each time fraction because of the short duration of each time fraction, and are uncorrelated among them and from time fraction to time fraction because the time interval between time fractions is much longer that the correlation time of polarization rotation.

We consider that the dominant noise source at the coherent receiver is the amplified spontaneous emission (ASE) noise generated by optical amplifiers. The ASE noise field in each polarization is modelled as additive white Gaussian noise [22].

C. Coherent Detection Receiver

A coherent receiver with homodyne synchronous detection employing phase and polarization diversities [25] is considered. The outputs from the homodyne receiver are analog-to-digital converted and processed by digital signal processing (DSP) circuits, restoring the complex amplitude of the signal. Next, the decoder accomplishes the symbol decoding. We consider a perfectly balanced optical front-end with the photodetectors modeled as square-law devices with a unit

responsivity and without frequency limitations. Following the analysis presented in [22], [25], the detected ICXT term in polarization p is given by $E_{ICXT,p}(t)E^*_{LO,p}$, where Z^* is the complex conjugate of Z and $E_{LO,p}$ is the local oscillator (LO) electrical field in polarization p, given by $E_{LO,p} = \sqrt{P_{LO,p}}$ where $P_{LO,p}$ is the LO average power in polarization p, with $P_{LO,x} = P_{LO,y}$. The two polarization components are processed by the DSP to equalize the GVD and random polarization rotation, and to perform RRC filtering. Perfect carrier-phase estimation is assumed. With perfect compensation of GVD and random polarization rotation, the detected ICXT in the two polarization directions after the DSP, at the sampling time instant t_s, can be written as

$$\begin{bmatrix} I_{ICXT,x}(t_s) \\ I_{ICXT,y}(t_s) \end{bmatrix} = \mathbf{K}^{-1} \begin{bmatrix} \{E_{ICXT,x}(t)E^*_{LO,x}\} * h_E(t) \\ \{E_{ICXT,y}(t)E^*_{LO,y}\} * h_E(t) \end{bmatrix}_{t=t_s} \tag{14}$$

where $h_E(t) = h_{RRC}(t) * \mathfrak{F}^{-1}\{e^{j\overline{\beta}_n(\omega)L}\}$ and $\mathbf{K}^{-1}$ is

$$\mathbf{K}^{-1} = \begin{bmatrix} e^{-j\theta}\cos(\Gamma) & e^{-j\psi}\sin(\Gamma) \\ -e^{j\psi}\sin(\Gamma) & e^{j\theta}\cos(\Gamma) \end{bmatrix} \tag{15}$$

and, for the RC pulse shape, $t_s = sT_s$ with $s \in \mathbb{Z}$. The impulse response $h_E(t)$ corresponds to the compensation of GVD and RRC filtering, and the matrix $\mathbf{K}^{-1}$ performs the compensation of the random polarization rotation.

III. Variance of the Detected ICXT

The influence of the ICXT on the performance of the WC-MCF with coherent detection can be assessed by analyzing the variance of the coherently detected ICXT at the decoder input [3], [17]. In Appendix A, it is shown that the four components of the detected ICXT at the decoder input and time instant t_s have equal variance given by

$$\sigma^2_{ICXT,p,u}(t_s) = \frac{X_c P_n P_{LO,p}}{4}\left(1 - \frac{\beta_r}{4}\right)\left[1 + \frac{\beta_r}{4-\beta_r}K_s\right] \tag{16a}$$

where β_r is the roll-off factor of the RC pulse, and the function K_s describes the dependence of the detected ICXT variance on the skew between cores and is given by

$$K_s = \operatorname{sinc}(S_{mn}R_s)\cos\left[\pi S_{mn}R_s - \frac{2\pi}{T_s}(t_s - \tau_{mn})\right] \tag{16b}$$

with $\operatorname{sinc}(x) = \sin(\pi x)/(\pi x)$, R_s the symbol rate, and the ICXT level, X_c, defined as

$$X_c = \frac{N|\overline{K}_{nm}|^2 P_m}{P_n} \tag{17}$$

with the average power of the ICXT given by $N|\overline{K}_{nm}|^2 P_m$.

Expressions (16a) and (16b) show that the compensation of the random polarization rotation does not affect the variance of the detected ICXT components. This is a result of the fact that compensation is implemented by a unitary matrix that does not unbalance the variance of each component.

Expression (16a) shows that, when $\beta_r = 0$, (i) the variance of all detected ICXT components does not depend on the skew between cores and $\sigma^2_{ICXT,p,u} = X_c P_n P_{LO,p}/4$, which corresponds to the coherent detection of the mean ICXT power per component; (ii) the variance of all detected ICXT components does not depend on the sampling time instant. In fact, in case of $\beta_r \neq 0$, the variance of the detected ICXT components is periodic along time with the same period as the data symbol period, meaning that, in each time fraction, the detected ICXT components are cyclostationary processes in the wide sense with period equal to the symbol period. In the particular case of $\beta_r = 0$, the detected ICXT components become stationary processes as a result of the band-limitation due to the RC pulse shape. Expression (16a) shows also that the skew-dependence of the variance of the detected ICXT components is ruled by skew×symbol rate. This dependence increases with the pulse shape roll-off factor, vanishing for $\beta_r = 0$. In addition, analysis of expression (16b) allows to conclude that K_s can range between 1 and -1 (the conditions to reach these values depend on the sampling time instant and relative time delay between signals in the two cores, besides the skew×symbol rate itself) meaning that the skew can affect the detected ICXT components variance only on a fraction $\pm\beta_r/(4-\beta_r)$ relative to the case with $K_s = 0$. Therefore, the largest impact of the skew occurs for $\beta_r = 1$ and is $\pm 1/3$.

We stress that $\sigma^2_{ICXT,p,u}(t_s)$, given by eq. (16a), is the variance of the u component of the detected ICXT in polarization p, which can be assessed from the detected ICXT observed over a very long time interval. At this point, we denote the variance of the component of the detected ICXT in each time fraction as $S^2_{ICXT,p,u}$, which we relate to the definition of sample variance. We notice that the mean value of $S^2_{ICXT,p,u}$ calculated over a large number of time fractions is given by expression (16a), since the mean value of the sample variance of a random variable is the population variance, as long as the samples are independent and identically distributed [26].

IV. Computation of BER and OSNR penalty

This section presents the analytical expressions that quantify the impact of the ICXT on the BER and the OSNR penalty.

To assess the impact of the ICXT on the BER, we need to know the statistics of the detected ICXT. Other works considered that, in the low coupling regime, the ICXT in MCFs may be modeled as additive white Gaussian noise [3], [17], [19]. We stress that, due to the dependence on the statistics of the signal in the interfering cores, it may be questionable that the statistics of the in-phase and quadrature components of the detected ICXT are also Gaussian-distributed. Analyses of the statistics of the detected ICXT components using numerical simulation allowed us to conclude that the statistics of each component of the detected ICXT is approximately Gaussian-distributed. This observation combined with the fact that the components of the detected ICXT are uncorrelated (as discussed in Appendix A) allows us to model the detected ICXT in each polarization approximately as virtual additive white Gaussian noise. Comparison of BER estimates obtained by numerical Monte-Carlo simulation with theoretical estimates of the BER presented in subsection V-B will provide further evidence of the validity of this approximation.

Because of this approximation for the detected ICXT statistics and its independence of ASE noise, the overall variance

of ASE noise and ICXT at the decoder input is the sum of the variances of the ASE noise and the detected ICXT at the decoder input, and the electrical signal-to-(noise+ICXT) ratio (SNR) at the decoder input may be written as

$$\text{SNR} = \frac{P_n P_{LO,p}}{\sigma^2_{ASE} + \sigma^2_{ICXT}} \tag{18a}$$

where σ^2_{ASE} is the total variance of the ASE noise of the two polarization directions, and σ^2_{ICXT} is the total variance of detected ICXT, given by

$$\sigma^2_{ICXT} = \sum_{p\in\{x,y\}} \sum_{u\in\{I,Q\}} \sigma^2_{ICXT,p,u}(t_s) \tag{18b}$$

Because the four detected ICXT components at the decoder input and time instant t_s have equal variance, we may write

$$\sigma^2_{ICXT} = 4\sigma^2_{ICXT,p,u}(t_s). \tag{19}$$

With RC pulse shape, there is no intersymbol interference at the decoder input. In this case, the theoretical BER of M-QAM format, BER, can be written as [27]

$$\text{BER} = 4\frac{(1-1/\sqrt{M})}{\log_2(M)} Q\left(\sqrt{\frac{3}{M-1}\text{SNR}}\right) \tag{20}$$

where $Q(x)$ is the Q function [27]. We stress that expression (20) gives the BER averaged over a long time of BER measurement, which may be different from the BER measured along a time fraction.

We denote the SNR in the absence of ICXT by SNR_0. We note that, from eq. (18a), the SNR and SNR_0 are related by

$$\text{SNR} = \left[\frac{1}{\text{SNR}_0} + \frac{\sigma^2_{ICXT}}{P_n P_{LO,p}}\right]^{-1}. \tag{21}$$

We define the OSNR penalty, ΔOSNR, as the ratio between the required OSNR to achieve a specific BER, in presence of ICXT, OSNR_{req}, and the required OSNR for the same BER in absence of ICXT, $\text{OSNR}_{\text{req},0}$. The required SNR to achieve a specific BER, SNR_{req}, can be obtained from eq. (20). By relating SNR_0 with the OSNR definition [22], we get

$$\text{OSNR}_{\text{req}} = \left[\frac{1}{\text{SNR}_{\text{req}}} - \frac{\sigma^2_{ICXT}}{P_n P_{LO,p}}\right]^{-1} \cdot \frac{R_b}{B_{\text{ref}}\log_2(M)} \tag{22}$$

where R_b is the bit rate and B_{ref} is the reference bandwidth, which is 12.5 GHz. The required OSNR in absence of ICXT, $\text{OSNR}_{\text{req},0}$, is obtained from eq. (22) by setting $\sigma^2_{ICXT} = 0$. Using eq. (22) and the definition of $\text{OSNR}_{\text{req},0}$, we obtain

$$\Delta\text{OSNR} = \left[1 - \frac{\sigma^2_{ICXT}}{P_n P_{LO,p}} \cdot \text{SNR}_{\text{req}}\right]^{-1}. \tag{23}$$

Substituting eqs. (19) and (16a) in eq. (23), we obtain the OSNR penalty as a function of ICXT parameters:

$$\Delta\text{OSNR} = \left[1 - X_c\left(1-\frac{\beta_r}{4}\right)\left(1+\frac{\beta_r}{4-\beta_r}K_s\right)\text{SNR}_{\text{req}}\right]^{-1}. \tag{24}$$

Expression (24) allows to estimate the OSNR degradation for arbitrary BER, skew between cores, modulation format order and roll-off factor. Expression (24) shows that, for a null roll-off factor, the OSNR penalty due to ICXT depends only on the ICXT level X_c. Expressions similar to expression (24) with $\beta_r = 0$ are reported in [16], [17], [20] but derived under different conditions. The increase of the roll-off factor augments the impact of the inter-core skew on the OSNR degradation due to the ICXT. Also, the use of modulation formats with higher order increases the required SNR and, hence, higher order modulation formats experience larger OSNR degradation due to ICXT, as already reported in [16]–[18]. A quantitative analysis of the impact of ICXT parameters on the OSNR penalty is presented in subsection V-C.

In the case of multiple interfering cores, the ICXT field at the coherent receiver input results from the sum of the ICXT field contributions associated with each one of the interfering cores [12], [15]. As these contributions are independent, because they are originated from independent data symbols transmitted in different cores and from independent RPSs associated with different interfering cores, the total variance of each detected ICXT component is the sum of the variances associated with each interfering core, being each variance given by eq. (16a) with the parameters involved in that expression corresponding to each interfering core. When all interfering cores have the same parameters, the total variance of each detected ICXT component can still be written as eq. (16a) with $X_c = N_i N|\overline{K}_{nm}|^2 P_m/P_n$ where N_i is the number of interfering cores. The expression of OSNR degradation due to the ICXT, expression (24), is correspondingly affected in presence of multiple interfering cores.

Table I: Parameters of the PDM-QAM WC-MCF system.

Parameter	Symbol	Value
Average power of core c	P_c	0 dBm
LO power	P_{LO}	3 dBm
Bit rate	R_b	112 Gbps
Roll-off factor	β_r	$0.01, 0.25, 1$
Time misalignment	τ_{mn}	0
Modulation order	M	4, 16, 64

V. Numerical Results and Discussion

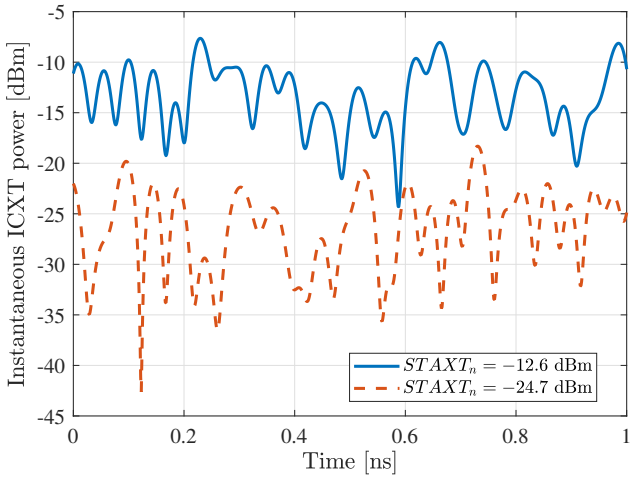
In this section, the influence of the inter-core skew on the STAXT and on the detected ICXT is analysed. Also, estimates from the analytical expressions for the average BER and OSNR penalty are compared with MC simulation results. Two cores are considered in all numerical results. Table I shows the value of the parameters used in our study. The numerical simulation results were obtained using the model presented in Section II and the theoretical (analytical) results were obtained using the expressions presented in Section IV.

A. STAXT and detected ICXT in each time fraction

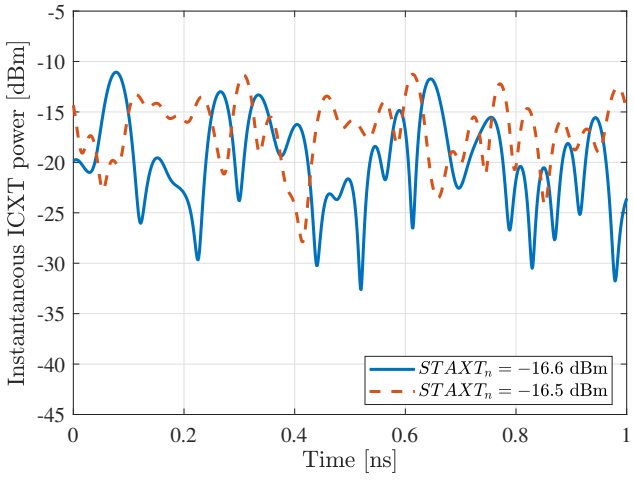
In this subsection, we analyse the impact of the inter-core skew on the STAXT and on the detected ICXT in each time fraction. In this analysis, the ICXT is induced by a PDM 4-QAM signal and $X_c = -16.7$ dB. This ICXT level leads to 1 dB OSNR penalty for a BER of 10^{-3}, a 4-QAM signal with $\beta_r = 0.01$ in the interfered core and $S_{mn}R_s = 0.01$, as will be shown in subsection V-C. This ICXT level was adopted as it represents the worst-case level corresponding to

the maximum acceptable OSNR degradation due to ICXT in WC-MCF transmission for that BER level. Other ICXT levels were analyzed and a similar behavior as the one reported in this subsection for $X_c = -16.7$ dB was observed. Each time fraction has a duration of 2^{18} PDM 4-QAM symbols (≈9.4 μs). The STAXT is calculated over the total duration of the time fraction. Further analyses showed that similar STAXT estimates are obtained using a duration of nearly 4.7 μs.

Figure 1 depicts the instantaneous ICXT power, given by expression (10), as a function of time, for two different time fractions, $\beta_r = 0.01$, and two cases of skew×symbol rate, $S_{mn}R_s = 0.01$ in Fig. 1-(a) and $S_{mn}R_s = 100$ in Fig. 1-(b). These two values of skew×symbol rate were chosen because they show the typical behavior for low and high skew×symbol rate. The two fractions were chosen arbitrarily but represent the typical evolution of ICXT power along time that occurs or low and high skew×symbol rate. Figure 1 shows that, for both high and low skew×symbol rate, the instantaneous ICXT power varies randomly with large fluctuations along time, that may exceed 20 dB peak-to-peak. In the following, we denote by $STAXT_n$ the STAXT level of each time fraction. Figure 1-(a) shows that, with reduced skew×symbol rate, different time fractions can have very distinct $STAXT_n$, -12.6 and -24.7 dBm in these time fractions. The large difference between these $STAXT_n$ indicates that the STAXT may change significantly over time when the ICXT is induced by a PDM QAM signal with symbol rate so that the skew×symbol rate is much lower than one, as also shown in [13]. For skew×symbol rate≫1, Fig. 1-(b), the difference between the $STAXT_n$ is only 0.1 dB, which is much smaller than the one observed in Fig. 1-(a), and the $STAXT_n$ is approximately equal to the average ICXT power.



(a)



(b)

Figure 1: Instantaneous ICXT power, in dBm, as a function of time with $X_c = -16.7$ dB and $\beta_r = 0.01$, for (a) $S_{mn}R_s = 0.01$ and (b) $S_{mn}R_s = 100$.

We define the normalized STAXT, $NSXT$, as the ratio between the STAXT and the average ICXT power. In the following, we use the mean and the variance of $NSXT$, $\mathbf{E}[NSXT]$ and σ^2_{NSXT} respectively, to quantify the STAXT fluctuations over time. Fig. 2 shows the mean and the variance of $NSXT$ as a function of skew×symbol rate, for $\beta_r = 0.01$ and $\beta_r = 1$. The numerical results in Fig. 2 are obtained using 20000 time fractions (each one with duration of ≈ 73 ns) to guarantee a good accuracy of the statistical estimates. It was confirmed that this shorter duration does not affect significantly the accuracy of the STAXT estimate and speeds up the computation of the statistical STAXT estimates. Figure 2 shows that the two values of β_r lead practically to the same mean and variance of STAXT. Figure 2 also shows that the mean value of the STAXT is independent of the skew×symbol rate and is equal to the average ICXT power. In addition, the variance of the normalized STAXT is 0.25 for $S_{mn}R_s < 0.1$. For higher values of skew×symbol rate, the normalized STAXT variance decreases, and is practically zero for $S_{mn}R_s > 10$. This dependence of STAXT on skew×symbol rate is in agreement with the results reported in [13], for $\beta_r = 0.01$. Results presented in Fig. 2 show that a similar dependence occurs for other values of β_r.

Results of Figs. 1 and 2 show that, although the variance of the STAXT may be practically zero, the instantaneous ICXT

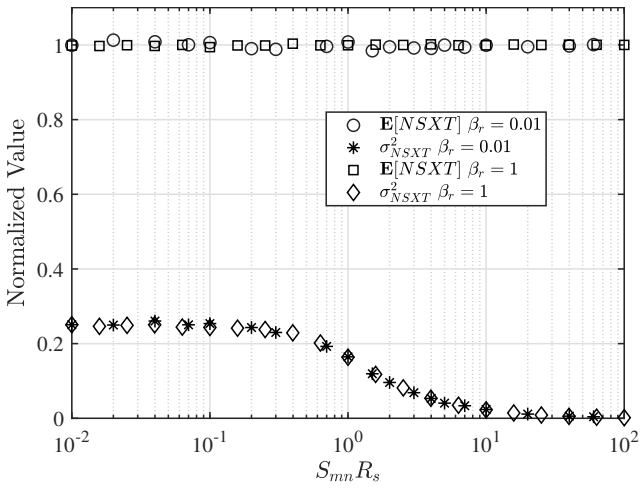


Figure 2: Mean (circles) and variance (asterisks) of $NSXT$ for $\beta_r = 0.01$ and mean (squares) and variance (diamonds) of $NSXT$ for $\beta_r = 1$ as a function of $S_{mn}R_s$.

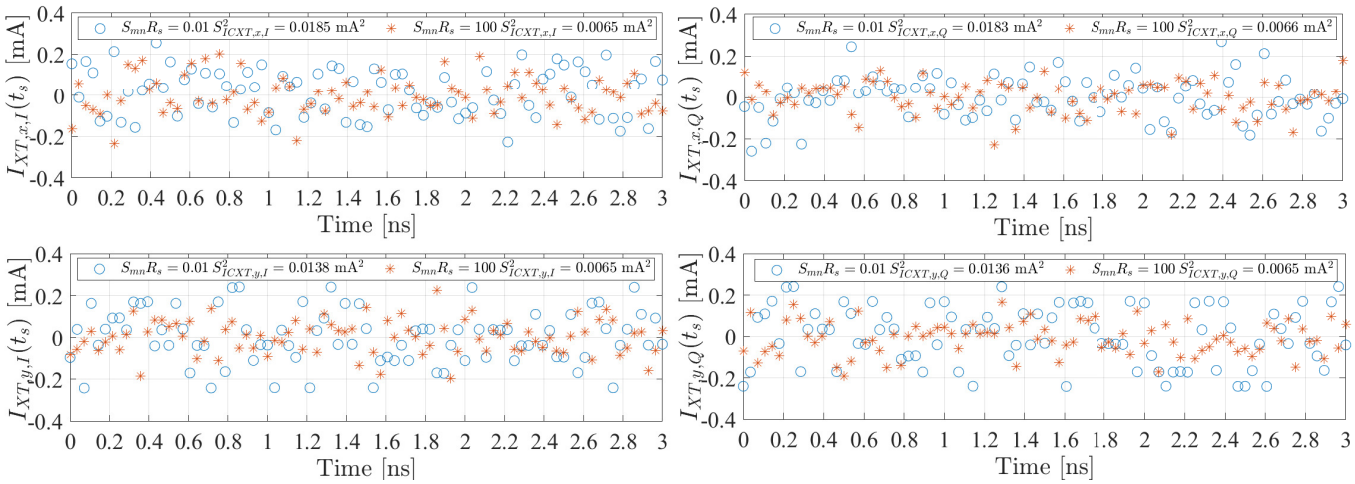


Figure 3: Components of the detected ICXT as a function of time, with $X_c = -16.7$ dB and $\beta_r = 0.01$, for (circle) $S_{mn}R_s = 0.01$ and (asterisk) $S_{mn}R_s = 100$.

power has very large fluctuations. In the following, we will analyse how these fluctuations affect the detected ICXT at the sampling time and the variance of each component of the detected ICXT in each time fraction. Figure 3 depicts the detected ICXT components at the sampling time instant as a function of time for a given time fraction. Figure 3 shows results for $S_{mn}R_s = 0.01$ and $S_{mn}R_s = 100$, with $\beta_r = 0.01$, as well as the corresponding variances calculated over the whole time fraction. The fluctuations of the detected ICXT observed in Fig. 3 along time for both skew×symbol rates are due to the fluctuations of the instantaneous ICXT power along time, as those depicted in Fig. 1. Results of Fig. 3 show that, with low skew×symbol rate, the variances of the I and Q components of the detected ICXT in the same polarization direction are approximately equal and, for different polarizations, can be slightly different. This behavior is also observed for the variances of the components of the detected ICXT with other roll-off factors, namely with $\beta_r = 1$. With high skew×symbol rate, the variances of the detected ICXT components and in both polarization directions are practically the same. These behaviors were observed also in other time fractions. In order to relate these behaviors of the variances of the detected ICXT components of each time fraction, $S^2_{ICXT,p,u}$, to the corresponding $STAXT_n$, we computed those quantities for 1000 time fractions, each one with a duration of 9.4 μs (2^{18} 4-QAM symbols). The results obtained for $S_{mn}R_s = 0.01$ and $S_{mn}R_s = 100$, and with $\beta_r = 0.01$, are depicted in the scatter plots of Fig. 4. Figure 4 shows that, for $S_{mn}R_s = 0.01$, the variance of the detected ICXT component may vary significantly from time fraction to time fraction. This variation is related to the level of STAXT, with an increase of spreading of values of $S^2_{ICXT,p,u}$ when the level of STAXT increases. For $S_{mn}R_s = 100$, the values of STAXT and variance of the detected ICXT component are concentrated near their mean value, and the variations of STAXT and variance of the detected ICXT components are much smaller than the ones observed for $S_{mn}R_s = 0.01$; in fact, they are almost zero. These results show that increasing the skew may reduce significantly the variation of the magnitude of the fluctuations (on the average) of each component of the detected ICXT from time fraction to time fraction ($S^2_{ICXT,p,u}$), being the magnitude of the fluctuations (on the average) similar in all time fractions when the skew is much larger than the symbol period. This is attributed to the fact that, when the skew is much larger than the symbol period, the ICXT power at each time instant results from the contribution of a large number of QAM symbols, each one with an independent random phase, causing that no significant differences are noticed in the STAXT and in the variance of each component of the detected ICXT (somehow related to a parcel of the STAXT) along time (from time fraction to time fraction). When the skew is much smaller than the symbol period, just one QAM symbol induces the ICXT along the MCF; thus, the effectiveness of the addition of all contributions of ICXT generated at the PMPs depends only on the RPSs occuring in the time fraction. In the time fractions where those RPSs are mostly in-phase, we have constructive addition and, on the average, the ICXT power increases; if they are out-of-phase, we have a reduction of the average ICXT power in the time fraction and, consequently, a reduction of $S^2_{ICXT,p,u}$.

We stress that the large fluctuations of the variance of the detected ICXT components from time faction to time fraction observed with $S_{mn}R_s = 0.01$ may cause significant BER fluctuations from time fraction to time fraction, cause sporadic BER events higher than a given BER threshold, leading to service interruption. To analyse the magnitude of the fluctuations of the detected ICXT variances of each time fraction and the impact of the inter-core skew and the roll-off factor on those variances, we define the normalized detected ICXT variance in polarization p and component u, $\widetilde{S}^2_{ICXT,p,u}$, as the ratio between the variance of the detected ICXT component in each time fraction and the total variance of the detected ICXT with zero roll-off factor, i. e., $\widetilde{S}^2_{ICXT,p,u} = S^2_{ICXT,p,u}/(X_c P_m P_{LO,p})$. Figure 5 shows the mean and the variance of $\widetilde{S}^2_{ICXT,p,u}$ as a function of $S_{mn}R_s$ for $\beta_r = 0.01$ and $\beta_r = 1$. The mean value is indicated as $\mathbf{E}[\widetilde{S}^2_{ICXT,p,u}]$,

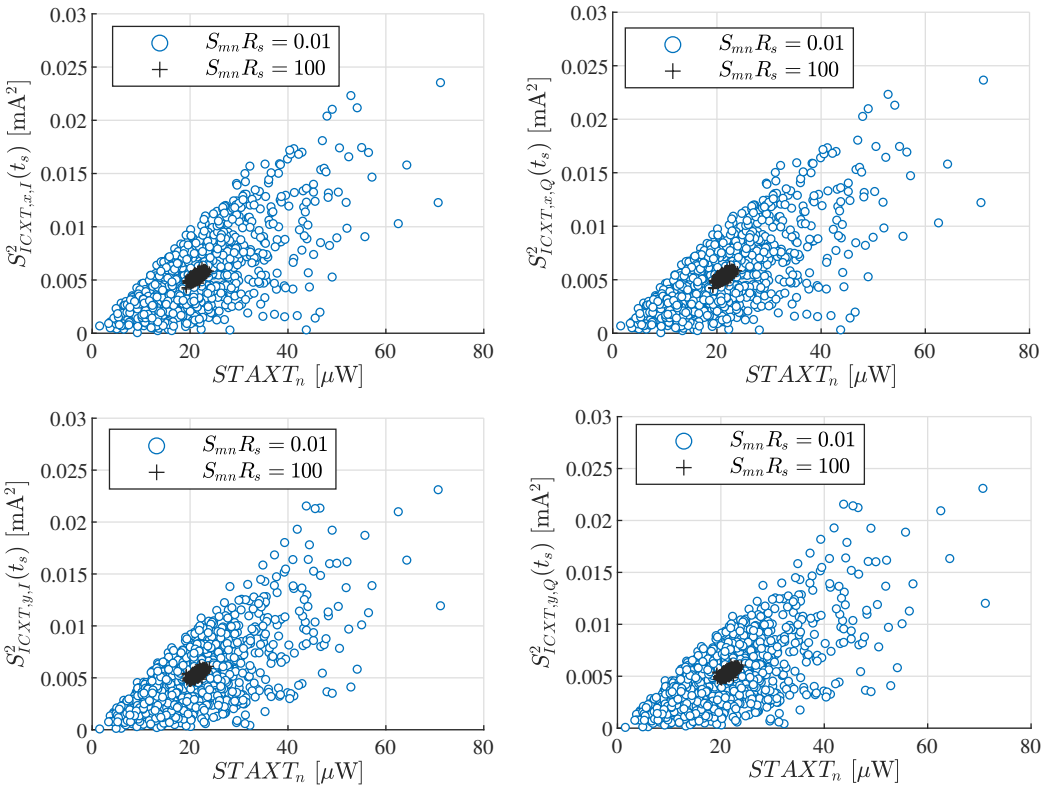


Figure 4: Variance of each component of the detected ICXT in each time fraction versus $STAXT_n$, with $X_c = -16.7$ dB and $\beta_r = 0.01$, for (circle) $S_{mn}R_s = 0.01$ and (plus sign) $S_{mn}R_s = 100$.

while the variance is denoted as Var($\widetilde{S}^2_{ICXT,p,u}$). The mean and the variance of $\widetilde{S}^2_{ICXT,p,u}$ are estimated from 2×10^4 time fractions, each one with 2^{11} 4-QAM symbols. For $\beta_r = 0.01$, Fig. 5 shows that the mean value of the variance of the detected ICXT component in each time fraction is independent of the skew. This is in agreement with expression (16a). For $\beta_r = 1$, Fig. 5 shows that the mean value of the detected ICXT variance of each time fraction depends on the skew×symbol rate, while the mean value of the STAXT is independent of this product, as shown in Figure 2. For very high skew×symbol rate, the mean value of the normalized variance of the component of the detected ICXT per time fraction tends to 0.1875, which is in agreement with expression (16a).

The variance of the normalized variance of the detected ICXT components in each time fraction provides a measure of the magnitude of the fluctuations of the variance from time fraction to time fraction. Fig. 5 shows that the increase of the skew×symbol rate product leads a remarkable reduction of the fluctuations of the detected ICXT variance between time fractions for both values of β_r, being those fluctuations practically zero for $S_{mn}R_s > 10$.

B. BER estimation

In this subsection, we analyze the influence of the ICXT on the BER of each time fraction. Also, the accuracy of the theoretical average BER obtained from expression (20) is assessed for different ICXT levels and different modulation orders, by comparison with the BER estimated by MC simulation.

Figure 6 depicts the scatter plot of the BER in each time fraction and the corresponding STAXT, for $S_{mn}R_s = 0.01$ and $S_{mn}R_s = 100$, for a BER of 10^{-3} in absence of ICXT. The time fractions in Fig. 6 are the same as those considered in Fig. 4, i. e. 1000 time fractions (each one with duration of $9.4\,\mu$s) of ICXT induced by a PDM 4-QAM interfering signal with ICXT level of -16.7 dB. We observe that, in this case of low ICXT level, for $S_{mn}R_s = 0.01$, the BERs can vary about one order of magnitude from time fraction to time fraction. Figure 6 shows that this variation may be associated with the large STAXT variation. For $S_{mn}R_s = 100$, the BERs in different time fractions are concentrated around $10^{-2.6}$ with a maximum logarithmic BER difference of 0.2, and the STAXT power varies only about ±2 μW around 21 μW.

The fluctuations of BER from time fraction to time fraction shown in the simulation results of Fig. 6 are attributed to the fluctuations of the detected ICXT variance from time fraction to time fraction, as these fluctuations lead to a variation of the BER in a similar manner. The large increase of the BER fluctuations for $S_{mn}R_s = 0.01$ relative to $S_{mn}R_s = 100$ allows to infer that WC-MCF systems with coherent detection and skew×symbol rate much lower than one are much more susceptible to suffer from outage periods than systems with skew×symbol rate much higher than one.

Figure 7 shows theoretical and numerical simulation results of the average BER (averaged along a number of time

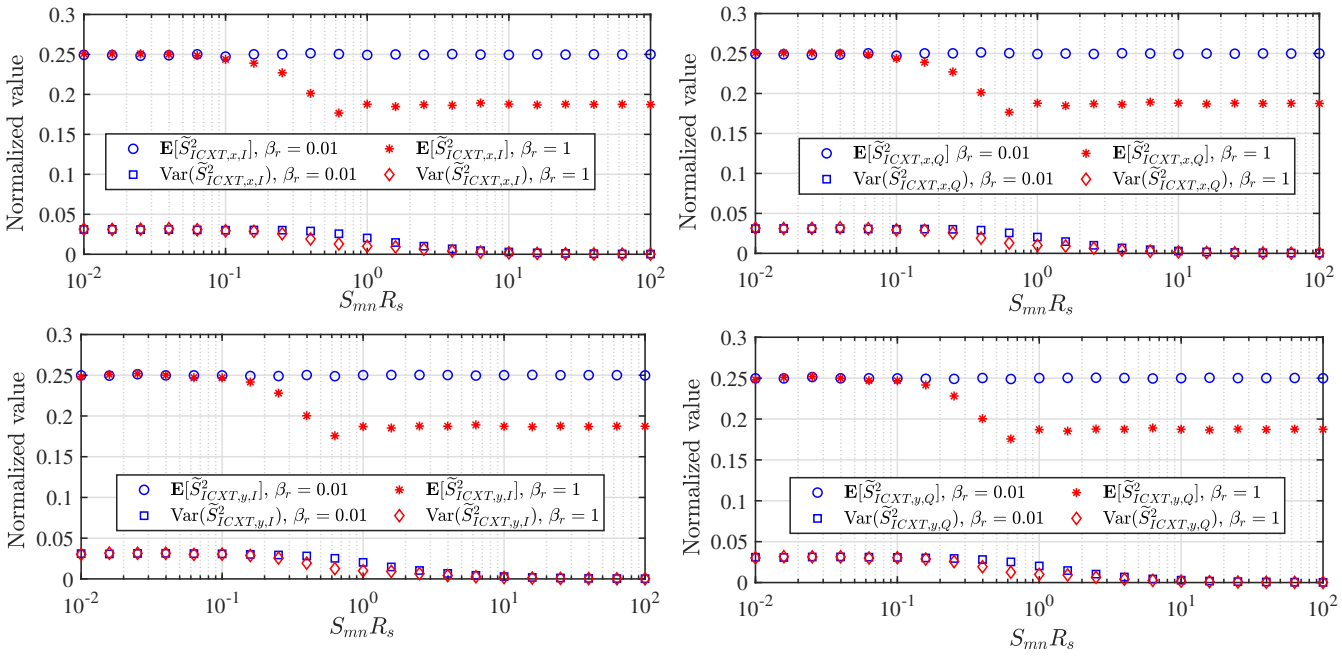


Figure 5: Mean value (circles) and variance of the normalized $S^2_{ICXT,p,u}$ (squares) with β_r of 0.01, and mean value (asterisks) and variance of the normalized $S^2_{ICXT,p,u}$ (diamonds) with β_r of 1 as a function of $S_{mn}R_s$.

fractions) as a function of OSNR, with $\beta_r = 0.01$ and $S_{mn}R_s = 0.01$. The dashed lines correspond to the theoretical average BER in absence of ICXT, and the solid lines to the theoretical average BER with ICXT levels of -16.7, -23.7 and -29.5 dB for the 4, 16 and 64-QAM formats, respectively. According to (24), these ICXT levels lead to 1 dB OSNR penalty for a BER of 10^{-3} and $S_{mn}R_s = 0.01$. The theoretical average BERs are obtained from expression (20). In Fig. 7, the symbols correspond to the BER obtained from MC simulation averaged along 500 time fractions, each one with 2^{11} transmitted symbols in each polarization direction and with stopping criterion of 4000 errored symbols. These figures allow to have a good accuracy in the estimation of the BER, for BERs of about 10^{-3} in presence of ICXT [21]. Figure 7 shows good agreement between the simulation and theoretical results. In particular, for each modulation format, the ICXT level imposed leads to 1 dB OSNR penalty for a BER of 10^{-3}, in the MC simulation and theoretically. The theoretical average BER given by (20) has been also compared with MC simulation results for higher skew×symbol rate, different roll-off factors, and 4, 16 and 64-QAM formats. A good agreement with the MC simulation results was noticed.

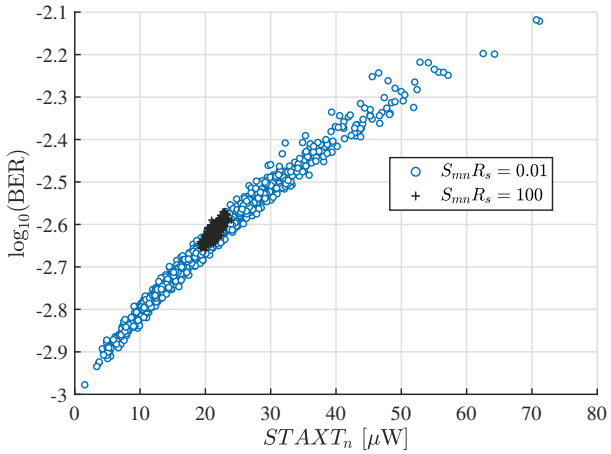


Figure 6: BER versus $STAXT_n$ with $X_c = -16.7$ dB and $\beta_r = 0.01$, for (circle) $S_{mn}R_s = 0.01$ and (plus) $S_{mn}R_s = 100$.

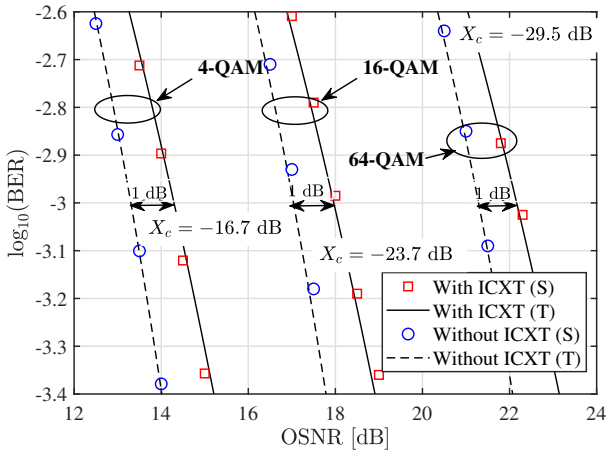


Figure 7: Average BER as a function of OSNR for different ICXT levels, with $S_{mn}R_s = 0.01$ and $\beta_r = 0.01$. (S): numerical simulation results; (T): theoretical results obtained using expression (20).

C. OSNR penalty

In this subsection, we assess the OSNR penalty due to ICXT using MC simulation and the analytical expression (24). Figure 8 depicts the OSNR penalty as a function of the ICXT level, for 4-QAM, 16-QAM and 64-QAM signals in both cores, $\beta_r = 0.01$, and $S_{mn}R_s = 0.01$. MC simulation and theoretical results are shown for the average BERs of 1.5×10^{-2} and 10^{-3}. Figure 8 shows very good agreement between the results of OSNR penalty estimated by MC simulation and theoretically for BER of $1.5{\times}10^{-2}$. For BER of 10^{-3}, Fig. 8 shows that the MC simulation and the theoretical results are in agreement for OSNR penalties below 1 dB. For higher OSNR penalties, the increase of the ICXT level leads to slight discrepancies between the MC simulation and the theoretical results that do not exceed 0.5 dB in the range of OSNR penalties tested. These discrepancies are higher for higher modulation orders and are attributed to the fact that the

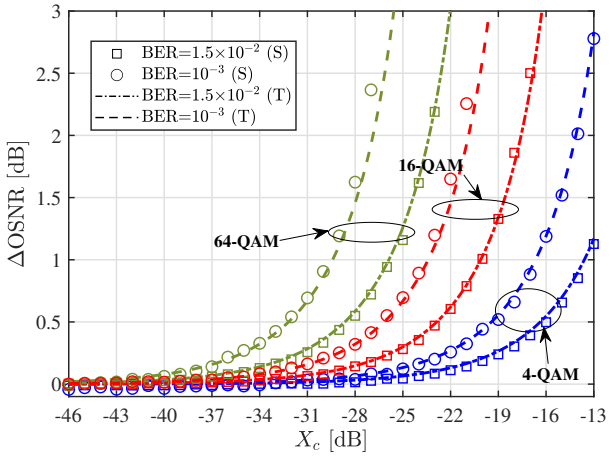


Figure 8: OSNR penalty as a function of the ICXT level for 4-QAM, 16-QAM and 64-QAM formats, with $\beta_r = 0.01$, $S_{mn}R_s = 0.01$ and different BERs. (S): simulation results; (T): theoretical results obtained using expression (24).

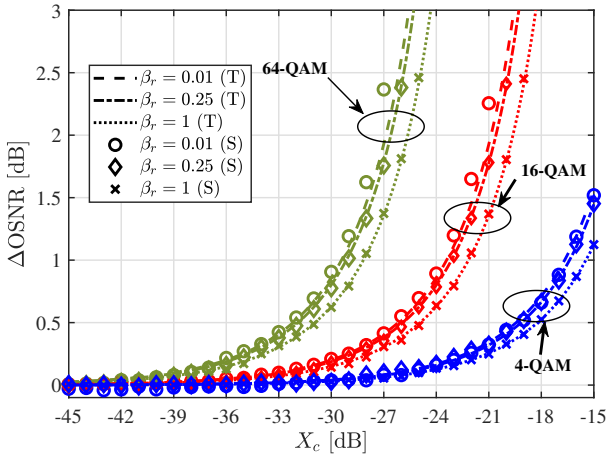


Figure 9: OSNR penalty as a function of the ICXT level for 4-QAM, 16-QAM and 64-QAM formats and β_r of 0.01, 0.25 and 1, with $S_{mn}R_s = 100$ and BER of 10^{-3}.

Table II: Maximum allowable ICXT levels of PDM M-QAM formats, in dB, for 1 dB OSNR penalty, for a BER of (A) 10^{-3} and (B) 1.5×10^{-2}, with $S_{mn}R_s = 100$.

		4-QAM		16-QAM		64-QAM	
		A	B	A	B	A	B
β_r	0.01	−16.7	−13.6	−23.4	−20.1	−29.5	−25.9
	0.25	−16.4	−13.3	−23.1	−19.8	−29.2	−25.6
	1	−15.4	−12.3	−22.1	−18.8	−28.2	−24.6

distribution of each component of the detected ICXT is farther from the Gaussian distribution when the modulation order increases. Figure 8 shows also that, for a given BER, lower modulation format orders increase the ICXT tolerance, as also reported in [16]–[18]. Figure 9 depicts the MC simulation and theoretical results of OSNR penalty for different M-QAM formats, β_r of 0.01, 0.25, and 1, $S_{mn}R_s = 100$ and a BER of 10^{-3}. Figure 9 shows that the agreement between the theoretical and MC simulation results of OSNR penalty improves with the increase of the roll-off factor and skew×symbol rate, even for OSNR penalties above 1 dB, with excellent agreement for β_r of 0.25 and 1. Figure 9 shows also that the increase of roll-off factor and skew enhances the ICXT tolerance. The results in Figs. 8 and 9 are in accordance with the results presented in Fig. 5 where, for $\beta_r = 1$, the ICXT variance is lower for higher values of skew×symbol rate.

The maximum allowable ICXT levels for 1 dB OSNR penalty, BERs of 1.5×10^{-2} and 10^{-3}, and $S_{mn}R_s = 100$ are presented in Table II. For $S_{mn}R_s = 0.01$, the maximum allowable ICXT level is the same for all roll-off factors, and it is the same as for $S_{mn}R_s = 100$ and $\beta_r = 0.01$. Table II shows that, for high skew-symbol rate product ($\gg$1) and $\beta_r = 0.25$, only a slight improvement of the ICXT tolerance by 0.3 dB for both BERs is noticed relative to $\beta_r = 0.01$. For $\beta_r = 1$ and both BERs, an improvement of the ICXT tolerance by about 1.3 dB relative to $\beta_r = 0.01$ is achieved.

It should be stressed out that, for nearly null roll-off factor, the results presented in Table II agree quite well with the theoretical ICXT levels presented in [16], [17]. For roll-off factor of nearly zero, those results agree quite well also with the crosstalk levels obtained experimentally in [18] for coherent QAM systems impaired by inband crosstalk, for the three modulation formats (4-, 16- and 64-QAM). We notice that the inband crosstalk was experimentally realized using optical couplers and variable optical attenuators [18]. This type of crosstalk causes BER degradation, which is constant along time, similarly to what occurs with ICXT for almost null roll-off factor and high skew×symbol rate, as can be seen in Fig. 6. We have also compared the OSNR penalty estimates obtained theoretically from expression (24) with the experimental results of SNR shown in Fig. 2 of [20] for the optimum launch power and the three system lengths. Using the values of the parameters of the experimental work, the ICXT level and the OSNR penalty for each system length have been obtained. Although the definition of SNR penalty considered in [20] is slightly different from the definition of OSNR penalty, the differences between the OSNR penalty and the experimental SNR penalty do not exceed 0.2 dB.

VI. Conclusion

We have analyzed the impact of the ICXT on the performance of WC-MCF systems with coherent detection and arbitrary inter-core skew. Our results show that, with low skew×symbol rate, the STAXT can have large fluctuations along the time and the detected ICXT variance can also have large fluctuations along the time leading to large fluctuations of BER along time. With large skew×symbol rate, although the STAXT may be almost constant along the time, the detected ICXT varies along time but its variance is almost constant along time, hence, leading to very small BER fluctuations along time.

To quantify the performance degradation due to ICXT for arbitrary skew, modulation format order and roll-off factor of the transmitted signals, analytical expressions for the variance of the coherently detected ICXT, average BER and OSNR penalty have been proposed and good agreement of their estimates with MC simulation results has been found. Numerical results have shown that, for most cases of interest corresponding to low roll-off factor, the improvement of the maximum allowable ICXT level for 1 dB OSNR penalty achieved by the increase of the skew is very small, and it is only 1.3 dB for a roll-off factor of 1. It has been also shown that the OSNR penalty estimates agree quite well with other authors' experimental results for nearly zero roll-off factor.

Appendix A
Variance and correlation properties of the detected ICXT components

Using eqs. (14) and (15), the detected ICXT in polarization x at the sampling time instant t_s at the decoder input can be written as

$$I_{ICXT,x}(t_s) = I_{x,x}(t_s) + I_{x,y}(t_s) \tag{25a}$$

where $I_{x,x}(t_s)$ is the detected ICXT in polarization x resulting from the x-component at the receiver input, given by

$$\begin{aligned} I_{x,x}(t_s) = &\frac{|\overline{K}_{nm}|\sqrt{P_m}\sqrt{P_{LO,x}}}{2}e^{-j\theta}\cos(\Gamma) \\ &\sum_{k=1}^{N}\sum_{i=-\infty}^{+\infty}[-ja_{x,i}^{(m)}e^{-j\Phi_{x,x}^{(k)}} - ja_{y,i}^{(m)}e^{-j\Phi_{y,x}^{(k)}}] \\ &h_{RC}(t_s - iT_s - \tau_{mn} - d_{mn}z_k) \end{aligned} \tag{25b}$$

and $I_{x,y}(t_s)$ is the detected ICXT in polarization x resulting from the y-component at the receiver input, given by

$$\begin{aligned} I_{x,y}(t_s) = &\frac{|\overline{K}_{nm}|\sqrt{P_m}\sqrt{P_{LO,y}}}{2}e^{-j\psi}\sin(\Gamma) \\ &\sum_{k=1}^{N}\sum_{i=-\infty}^{+\infty}[-ja_{x,i}^{(m)}e^{-j\Phi_{x,y}^{(k)}} - ja_{y,i}^{(m)}e^{-j\Phi_{y,y}^{(k)}}] \\ &h_{RC}(t_s - iT_s - \tau_{mn} - d_{mn}z_k). \end{aligned} \tag{25c}$$

Also using eqs. (14) and (15), the detected ICXT in polarization y at the sampling time instant t_s can be written as

$$I_{ICXT,y}(t_s) = I_{y,x}(t_s) + I_{y,y}(t_s) \tag{26a}$$

where $I_{y,x}(t_s)$ is the detected ICXT in polarization y resulting from the x-component at the receiver input, given by

$$\begin{aligned} I_{y,x}(t_s) = &-\frac{|\overline{K}_{nm}|\sqrt{P_m}\sqrt{P_{LO,x}}}{2}e^{j\psi}\sin(\Gamma) \\ &\sum_{k=1}^{N}\sum_{i=-\infty}^{+\infty}[-ja_{x,i}^{(m)}e^{-j\Phi_{x,x}^{(k)}} - ja_{y,i}^{(m)}e^{-j\Phi_{y,x}^{(k)}}] \\ &h_{RC}(t_s - iT_s - \tau_{mn} - d_{mn}z_k) \end{aligned} \tag{26b}$$

and $I_{y,y}(t_s)$ is the detected ICXT in polarization y resulting from the y-component at the receiver input, given by

$$\begin{aligned} I_{y,y}(t_s) = &\frac{|\overline{K}_{nm}|\sqrt{P_m}\sqrt{P_{LO,y}}}{2}e^{j\theta}\cos(\Gamma) \\ &\sum_{k=1}^{N}\sum_{i=-\infty}^{+\infty}[-ja_{x,i}^{(m)}e^{-j\Phi_{x,y}^{(k)}} - ja_{y,i}^{(m)}e^{-j\Phi_{y,y}^{(k)}}] \\ &h_{RC}(t_s - iT_s - \tau_{mn} - d_{mn}z_k). \end{aligned} \tag{26c}$$

We note that, for the four components, the GVD is perfectly compensated, and the detected ICXT has RC pulse shape as a consequence of $h_{RRC}(t) * h_{RRC}(t) = h_{RC}(t)$.

The in-phase and quadrature components in each polarization are obtained from eqs. (25) and (26) by extracting the real and imaginary parts, respectively. As the RPSs are uniformly distributed between 0 and 2π, the means of $I_{x,x}(t_s)$, $I_{x,y}(t_s)$, $I_{y,x}(t_s)$ and $I_{y,y}(t_s)$ are zero and, consequently, the mean of the detected ICXT of the in-phase and quadrature component in each polarization is zero.

As $I_{x,x}(t_s)$ and $I_{x,y}(t_s)$ are uncorrelated because they depend on statistically independent RPSs, the variance of the in-phase and quadrature components of polarization x (a similar analysis applies to polarization y) is given by

$$\begin{aligned} \mathbf{E}\big[I_{ICXT,x,I}(t_s)^2\big] = &\mathbf{E}\big[\,(\Re[I_{x,x}(t_s)])^2\,\big] \\ &+\mathbf{E}\big[\,(\Re[I_{x,y}(t_s)])^2\,\big] \end{aligned} \tag{27a}$$

and

$$\begin{aligned} \mathbf{E}\big[I_{ICXT,x,Q}(t_s)^2\big] = &\mathbf{E}\big[\,(\Im[I_{x,x}(t_s)])^2\,\big] \\ &+\mathbf{E}\big[\,(\Im[I_{x,y}(t_s)])^2\,\big] \end{aligned} \tag{27b}$$

For the I component, we have

$$\begin{aligned} &\mathbf{E}\big[\,(\Re[I_{x,x}(t_s)])^2\,\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,x}}{4}\cos^2(\Gamma) \\ &\sum_{k_1=1}^{N}\sum_{i_1=-\infty}^{+\infty}\sum_{k_2=1}^{N}\sum_{i_2=-\infty}^{+\infty}\mathbf{E}\Big[a_{I,x,i_1}^{(m)}a_{I,x,i_2}^{(m)}\sin\Phi_{x,x}^{(k_1)'}\sin\Phi_{x,x}^{(k_2)'} \\ &+ a_{Q,x,i_1}^{(m)}a_{Q,x,i_2}^{(m)}\cos\Phi_{x,x}^{(k_1)'}\cos\Phi_{x,x}^{(k_2)'} + a_{I,y,i_1}^{(m)}a_{I,y,i_2}^{(m)}\cdot \\ &\sin\Phi_{y,x}^{(k_1)'}\sin\Phi_{y,x}^{(k_2)'} + a_{Q,y,i_1}^{(m)}a_{Q,y,i_2}^{(m)}\cos\Phi_{y,x}^{(k_1)'}\cos\Phi_{y,x}^{(k_2)'}\Big] \\ &h_{RC}(t_s - i_1T_s - \tau_{mn} - d_{mn}z_{k_1})\cdot \\ &\qquad h_{RC}(t_s - i_2T_s - \tau_{mn} - d_{mn}z_{k_2}) \end{aligned} \tag{28a}$$

$$\mathbf{E}\big[\,(\Re[I_{x,y}(t_s)])^2\,\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,y}}{4}\sin^2(\Gamma)$$
$$\sum_{k_1=1}^{N}\sum_{i_1=-\infty}^{+\infty}\sum_{k_2=1}^{N}\sum_{i_2=-\infty}^{+\infty}\mathbf{E}\Big[a_{I,x,i_1}^{(m)}a_{I,x,i_2}^{(m)}\sin\Phi_{x,y}^{(k_1)'}\sin\Phi_{x,y}^{(k_2)'}$$
$$+\,a_{Q,x,i_1}^{(m)}a_{Q,x,i_2}^{(m)}\cos\Phi_{x,y}^{(k_1)'}\cos\Phi_{x,y}^{(k_2)'}+a_{I,y,i_1}^{(m)}a_{I,y,i_2}^{(m)}\cdot$$
$$\sin\Phi_{y,y}^{(k_1)'}\sin\Phi_{y,y}^{(k_2)'}+a_{Q,y,i_1}^{(m)}a_{Q,y,i_2}^{(m)}\cos\Phi_{y,y}^{(k_1)'}\cos\Phi_{y,y}^{(k_2)'}\Big]$$
$$h_{RC}(t_s-i_1T_s-\tau_{mn}-d_{mn}z_{k_1})\cdot$$
$$h_{RC}(t_s-i_2T_s-\tau_{mn}-d_{mn}z_{k_2}) \tag{28b}$$

in which the expected values of the cross terms between the amplitudes of the I and Q components of the transmitted symbols and of the cross terms between the amplitudes of the transmitted symbols in different polarizations were omitted because they are zero. The phases in eqs. (28) are defined as $\Phi_{p,x}^{(k)'}=\Phi_{p,x}^{(k)}+\theta$ and $\Phi_{p,y}^{(k)'}=\Phi_{p,y}^{(k)}+\psi$. Taking into account that $\mathbf{E}\left[\left(a_{I,p,i}^{(m)}\right)^2\right]=\mathbf{E}\left[\left(a_{Q,p,i}^{(m)}\right)^2\right]=T_s^2/2$ and

$$\mathbf{E}[\cos(\Phi_{b,p}^{(k_1)'})\cos(\Phi_{b,p}^{(k_2)'})]$$
$$=\mathbf{E}[\sin(\Phi_{b,p}^{(k_1)'})\sin(\Phi_{b,p}^{(k_2)'})]=\begin{cases}0, & k_1\neq k_2\\ \frac{1}{2} & k_1=k_2\end{cases} \tag{29}$$

the variance of $\Re[I_{x,x}(t_s)]$ and the variance of $\Re[I_{x,y}(t_s)]$ can be written, respectively, as

$$\mathbf{E}\big[\,(\Re[I_{x,x}(t_s)])^2\,\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,x} T_s^2}{4}\cos^2(\Gamma)$$
$$\sum_{i=-\infty}^{+\infty}\sum_{k=1}^{N}h_{RC}^2(t_s-iT_s-\tau_{mn}-d_{mn}z_k) \tag{30a}$$

$$\mathbf{E}\big[\,(\Re[I_{x,y}(t_s)])^2\,\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,y} T_s^2}{4}\sin^2(\Gamma)$$
$$\sum_{i=-\infty}^{+\infty}\sum_{k=1}^{N}h_{RC}^2(t_s-iT_s-\tau_{mn}-d_{mn}z_k). \tag{30b}$$

Thus, following expression (27a), we obtain

$$\mathbf{E}\big[I_{ICXT,x,I}(t_s)^2\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,p} T_s^2}{4}$$
$$\sum_{i=-\infty}^{+\infty}\sum_{k=1}^{N}h_{RC}^2(t_s-iT_s-\tau_{mn}-d_{mn}z_k) \tag{31}$$

where $P_{LO,x}=P_{LO,y}=P_{LO,p}$.

Following a similar analysis, the variance of the Q component of the detected ICXT of polarization x is

$$\mathbf{E}\big[I_{ICXT,x,Q}(t_s)^2\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,p} T_s^2}{4}$$
$$\sum_{i=-\infty}^{+\infty}\sum_{k=1}^{N}h_{RC}^2(t_s-iT_s-\tau_{mn}-d_{mn}z_k) \tag{32}$$

We note that the variance of the Q component of polarization x is equal to the variance of the I component, given by expression (31).

The I and Q components of the detected ICXT of polarization y are given, respectively, by

$$\mathbf{E}\big[I_{ICXT,y,I}(t_s)^2\big] = \mathbf{E}\big[\,(\Re[I_{y,x}(t_s)])^2\,\big] + \mathbf{E}\big[\,(\Re[I_{y,y}(t_s)])^2\,\big] \tag{33a}$$

$$\mathbf{E}\big[I_{ICXT,y,Q}(t_s)^2\big] = \mathbf{E}\big[\,(\Im[I_{y,x}(t_s)])^2\,\big] + \mathbf{E}\big[\,(\Im[I_{y,y}(t_s)])^2\,\big] \tag{33b}$$

as $I_{y,x}(t_s)$ and $I_{y,y}(t_s)$ are uncorrelated because they depend on statistically independent RPSs.

Following the same calculation steps as performed for the polarization x, we obtain for the variance of the I component of the detected ICXT of polarization y,

$$\mathbf{E}\big[I_{ICXT,y,I}(t_s)^2\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,p} T_s^2}{4}$$
$$\sum_{i=-\infty}^{+\infty}\sum_{k=1}^{N}h_{RC}^2(t_s-iT_s-\tau_{mn}-d_{mn}z_k) \tag{34}$$

and for the Q component

$$\mathbf{E}\big[I_{ICXT,y,Q}(t_s)^2\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,p} T_s^2}{4}$$
$$\sum_{i=-\infty}^{+\infty}\sum_{k=1}^{N}h_{RC}^2(t_s-iT_s-\tau_{mn}-d_{mn}z_k) \tag{35}$$

We note that the variances of the I and Q components of polarization y, expressions (34) and (35), are equal to the variances of the I and Q components of polarization x, given by expressions (31) and (32), respectively.

To simplify the sums in expressions (31), (32), (34) and (35), we use the Poisson summation formula [27]. Considering the variance of the I component of polarization x (for the other three components, the same simplification can be applied to), we obtain

$$\mathbf{E}\big[I_{ICXT,x,I}(t_s)^2\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,p} T_s}{4}$$
$$\sum_{k=1}^{N}\sum_{\nu=-\infty}^{+\infty}G\left(\frac{\nu}{T_s}\right)\cdot e^{j2\pi\frac{\nu}{T_s}(t_s-\tau_{mn}-d_{mn}z_k)} \tag{36}$$

where $G(f)=H_{RC}(f)*H_{RC}(f)$ with $H_{RC}(f)$ defined as

$$H_{RC}(f)=\begin{cases}1 & 0\le|f|\le\frac{1-\beta_r}{2T_s}\\ \cos^2\left[\frac{\pi T_s}{2\beta_r}\left(|f|-\frac{1-\beta_r}{2T_s}\right)\right] & \frac{1-\beta_r}{2T_s}\le|f|\le\frac{1+\beta_r}{2T_s}\\ 0 & |f|>\frac{1+\beta_r}{2T_s}\end{cases} \tag{37}$$

where β_r is the roll-off factor of the raised-cosine pulse shape. The function $G(f)$ holds:

$$G\left(\frac{\nu}{T_s}\right)=\begin{cases}\frac{1}{T_s}(1-\frac{\beta_r}{4}), & \nu=0\\ \frac{1}{T_s}\frac{\beta_r}{8}, & \nu=\pm1\\ 0, & \text{otherwise}\end{cases}. \tag{38}$$

Using eq. (38) in eq. (36), approximating $z_k\approx k\Delta z$, and taking into account that

$$\sum_{k=1}^{N}e^{-j2\pi\frac{\nu}{T_s}d_{mn}k\Delta z}=e^{-j2\pi\frac{\nu}{T_s}d_{mn}\Delta z}\left[\frac{1-e^{-j2\pi\frac{\nu}{T_s}d_{mn}\Delta zN}}{1-e^{-j2\pi\frac{\nu}{T_s}d_{mn}\Delta z}}\right], \tag{39}$$

we obtain

$$\mathbf{E}\big[|I_{ICXT,x,I}(t_s)|^2\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,p} T_s}{4}\Bigg[NG(0)+ 2G\left(\frac{1}{T_s}\right)\cos\left(\frac{\pi}{T_s}d_{mn}\Delta z(N+1)-\frac{2\pi}{T_s}(t_s-\tau_{mn})\right) \cdot \frac{\sin\left(\frac{\pi}{T_s}d_{mn}\Delta z N\right)}{\sin\left(\frac{\pi}{T_s}d_{mn}\Delta z\right)}\Bigg] \quad (40)$$

In the limit of a very large number of phase matching points ($N \to +\infty$), which corresponds to $\Delta z \to 0$, so that $N\Delta z = L$, with L the MCF length, eq. (40) is written as

$$\mathbf{E}\big[I_{ICXT,x,I}(t_s)^2\big] = \frac{|\overline{K}_{nm}|^2 P_m P_{LO,p} T_s}{4}\cdot \Bigg[NG(0) + 2NG\left(\frac{1}{T_s}\right)\cdot \operatorname{sinc}\left(S_{mn}R_s\right)\cdot \cos\left(\pi S_{mn}R_s - \frac{2\pi}{T_s}(t_s-\tau_{mn})\right)\Bigg]. \quad (41)$$

where $R_s = 1/T_s$. Substituting (38) in (41), we obtain

$$\mathbf{E}\big[|I_{ICXT,x,I}(t_s)|^2\big] = \frac{X_c P_n P_{LO,p}}{4}\left(1-\frac{\beta_r}{4}\right)\Bigg[1+ \frac{\beta_r}{4-\beta_r}\operatorname{sinc}\left(S_{mn}R_s\right)\cos\left(\pi S_{mn}R_s - \frac{2\pi}{T_s}(t_s-\tau_{mn})\right)\Bigg] \quad (42)$$

We stress that expression (42) also holds for the variance of the Q component of the detected ICXT of polarization x and of the I and Q components of polarization y. In the main text, we denote the variance of each one of these four components as $\sigma^2_{ICXT,p,u}(t_s)$.

The computation of the correlation between the I and Q components of polarizations x and y of the detected ICXT follows similar calculation steps to those performed in the computation of the variance of each component. Doing so, we conclude that the correlation between different components is zero for all combinations of components.

REFERENCES

[1] W. Klaus, B. Putnam, R. Luís, J. Sakaguchi, J. Mendinueta, Y. Awaji, and N. Wada, "Advanced space division multiplexing technologies for optical networks," *IEEE J. Opt. Commun. Netw.*, vol. 9, no. 4, pp. C1–C11, Apr. 2017.

[2] B. Puttnam, R. Luís, G. Rademacher, A. Alfredsson, W. Klaus, J. Sakaguchi, Y. Awaji, E. Agrell, and N. Wada, "Characteristics of homogeneous multi-core fibers for SDM transmission," *APL Photon. J.*, vol. 4, no. 2, Feb. 2019, art. no. 022804.

[3] J. Gené and P. Winzer, "A universal specification for multicore fiber crosstalk," *IEEE Photon. Technol. Lett.*, vol. 31, no. 9, pp. 673–676, May 1, 2019.

[4] T. Mizuno, H. Takara, K. Shibahara, A. Sano, and Y. Miyamoto, "Dense space division multiplexed transmission over multicore and multimode fiber for long-haul transport systems," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1484–1493, Mar. 15, 2016.

[5] B. Puttnam, G. Rademacher, R. Luís, T. Eriksson, W. Klaus, Y. Awaji, N. Wada, K. Maeda, S. Takasaka, and R. Sugizaki, "High data-rate and long distance MCF transmission with 19-core C+L band cladding-pumped EDFA," *J. Lightw. Technol.*, vol. 38, no. 1, pp. 123–130, Jan. 1, 2020.

[6] R. Luís, B. Puttnam, G. Rademacher, Y. Awaji, and N. Wada, "Crosstalk impact on a 535 Tb/s 172 km transmission using a homogeneous 19-core multicore fiber," in *Conference on Lasers and Electro-Optics (CLEO)*, Aug. 2018, paper SM4C.1.

[7] M. Koshiba, K. Saitoh, K. Takenaga, and S. Matsuo, "Analytical expression of average power-coupling coefficients for estimating intercore crosstalk in multicore fibers," *IEEE Photon. J.*, vol. 4, no. 5, pp. 1987–1995, Oct. 2012.

[8] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, "Design and fabrication of ultra-low crosstalk and low-loss multi-core fiber," *Opt. Express*, vol. 19, no. 17, pp. 16 576–16 592, Aug. 15, 2011.

[9] A. Cartaxo and T. Alves, "Discrete changes model of inter-core crosstalk of real homogeneous multi-core fibers," *J. Lightw. Technol.*, vol. 35, no. 12, pp. 2398–2408, Jun. 15, 2017.

[10] R. Soeiro, T. Alves, and A. Cartaxo, "Dual polarization discrete changes model of inter-core crosstalk in multi-core fibers," *IEEE Photon. Technol. Lett.*, vol. 29, no. 16, pp. 1395–1398, Aug. 15, 2017.

[11] R. Luís, B. Puttnam, A. Cartaxo, W. Klaus, J. Mendinueta, Y. Awaji, N. Wada, T. Nakanishi, T. Hayashi, and T. Sasaki, "Time and modulation frequency dependence of crosstalk in homogeneous multi-core fibers," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 441–447, Jan. 15, 2016.

[12] T. Alves and A. Cartaxo, "Characterization of the stochastic time evolution of short-term average intercore crosstalk in multicore fibers with multiple interfering cores," *Opt. Express*, vol. 26, no. 4, pp. 4605–4620, Feb. 2018.

[13] G. Rademacher, R. Luís, B. Puttnam, Y. Awaji, and N. Wada, "Crosstalk dynamics in multi-core fibers," *Opt. Express*, vol. 25, no. 10, pp. 12 020–12 028, May 15, 2017.

[14] A. Cartaxo, R. Luís, B. Puttnam, T. Hayashi, Y. Awaji, and N. Wada, "Dispersion impact on the crosstalk amplitude response of homogeneous multi-core fibers," *IEEE Photon. Technol. Lett.*, vol. 28, no. 17, pp. 1858–1861, Sep. 1, 2016.

[15] T. Alves and A. Cartaxo, "Decorrelation bandwidth of intercore crosstalk in weakly coupled multicore fibers with multiple interfering cores," *J. Lightw. Technol.*, vol. 37, no. 3, pp. 744–754, Feb. 1, 2019.

[16] B. Puttnam, R. Luís, T. Eriksson, W. Klaus, J. Mendinueta, Y. Awaji, and N. Wada, "Impact of inter-core crosstalk on the transmission distance of QAM formats in multi-core fibers," *IEEE Photon. J.*, vol. 8, no. 2, Apr. 2016, art. no. 0601109.

[17] T. Hayashi, T. Sasaki, and E. Sasaoka, "Behavior of inter-core crosstalk as a noise and its effect on Q-factor in multi-core fiber," *IEICE Transactions on Communications*, vol. E97-B, no. 5, pp. 936–944, May 2014.

[18] P. Winzer, A. Gnauck, A. Konczykowska, F. Jorge, and J.-Y. Dupuy, "Penalties from in-band crosstalk for advanced optical modulation formats," in *37th European Conference and Exhibition on Optical Communication (ECOC)*, Sep. 2011, paper Tu.5.B.7.

[19] T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, and E. Sasaoka, "Uncoupled multi-core fiber enhancing signal-to-noise ratio," *Opt. Express*, vol. 20, no. 26, pp. B94–B103, Dec 2012.

[20] R. Luís, G. Rademacher, B. Puttnam, D. Semrau, R. Killey, P. Bayvel, Y. Awaji, and H. Furukawa, "Crosstalk impact on the performance of wideband multicore-fiber transmission systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 26, no. 4, Jul. 2020, art. no. 4400609.

[21] B. Pinheiro, J. Rebola, and A. Cartaxo, "Inter-core crosstalk dependence on design parameters in coherent detection weakly-coupled multicore fiber systems," in *20th International Conference on Transparent Optical Networks (ICTON)*, Jul. 2018, paper Tu.D1.6.

[22] R. Essiambre, G. Kramer, P. Winzer, G. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 15, 2010.

[23] G. Bosco, "Spectrally efficient multiplexing: Nyquist-WDM," in *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks*. John Wiley & Sons, Ltd, 2016, ch. 4, pp. 123–156.

[24] P. Poggiolini and Y. Jiang, "Recent advances in the modeling of the impact of nonlinear fiber propagation effects on uncompensated coherent transmission systems," *J. Lightw. Technol.*, vol. 35, no. 3, pp. 458–480, Feb. 1, 2017.

[25] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 157–179, Jan. 1, 2016.

[26] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill, 1991.

[27] A. Carlson and P. Crilly, *Communication Systems: an introduction to signals and noise in electrical communication*, 5th ed. McGraw-Hill Education, 2009.