Joint Channel Estimation, Interference Mitigation, and Decoding for WDM Coherent Optical Communications

Frederic Lehmann, Petros Ramantanis, and Yann Frignac

***Abstract*—High spectral efficiency in wavelength-division multiplexing (WDM) coherent optical communications can be achieved by lowering the channel spacing to a value close to the baud rate. As a result, such communication systems are subject to a high level of two-dimensional (2D) interference, originating from both chromatic dispersion (CD) within WDM channels and cross talk between WDM channels. We propose to digitally restore the transmitted information at the receiver side by jointly processing the WDM channels, using 2D interference cancellation methods interacting with a forward error correction decoder. Such schemes have been investigated extensively for known channel parameters; however, little is known about the error performance in the absence of channel state information. This issue is important, since in practice uncertainties about the physical parameters of the optical fiber and noise levels are unavoidable. In this paper, we propose optical channel parameter estimation schemes with limited impact on the overall complexity. First, a decision-aided carrier phase recovery, taking the presence of 2D interference into account, is designed to compensate the laser beat linewidth. Then we provide a technique to estimate the linear channel impairments for each WDM channel, based on a limited number of training symbols. The validity of the proposed method is evaluated by numerical simulations taking into account linear and nonlinear fiber impairments. In particular, we show that for high-order modulations, joint processing of spectrally overlapping WDM channels becomes mandatory to obtain satisfactory performances.**

***Index Terms*—Carrier phase recovery; Code-aided processing; Interference mitigation; Optical coherent communications; WDM; 2D channel estimation.**

I. Introduction

During the last decade, there has been a renewed interest in coherent optical communications [1,2]. High-speed digital signal processing (DSP) circuits already offer the possibility to increase the spectral efficiency by using higher order modulation formats and to compensate for transmission impairments in an efficient way. However, due to the increase of global Internet traffic, continual improvements of the data rates on optical fibers are needed. Coherent optical transmissions also enable the use of state-of-the-art forward error correction (FEC) codes, such as low-density parity-check (LDPC) [3] or product codes [4], along with soft-input soft-output (SISO) decoders. The combination of coherent detection and capacity-achieving codes is already widespread in wireline and wireless applications [5].

The focus of this paper is on long-haul wavelength-division multiplexing (WDM) coherent optical transmission over transoceanic or transcontinental distances. Whereas in a context of optical transparent networks, the different wavelengths of a WDM transmission propagating on a given link may come from different transmitters or be multiplexed using an add and drop multiplexer in a transparent node, we consider here the possibility of transmitting a WDM subband within which all channels experience the same propagation history. This may be, for example, analogous to the case of subcarrier signals in an orthogonal frequency-division multiplexed superchannel.

In installed transmission systems, the time delay between WDM channels is generally unknown; therefore we consider asynchronous WDM transmissions. For the sake of simplicity, polarization multiplexing (PolMux) is not considered here, and we ignore the polarization mode dispersion (PMD) in the fiber propagation model. Indeed, when performing numerical simulations taking into account Kerr nonlinearity, we will not assess the polarization cross-phase modulation. The considered linear fiber impairments include i) intersymbol interference (ISI), originating from imperfect chromatic dispersion (CD) compensation and low-pass filtering at the receiver, and ii) cross talk between WDM channels. We will take into account phase noise due to the laser linewidth.

It is well known that an effective way to augment the capacity is to multiplex a number of data streams on different wavelengths (WDM) [2]. To further increase the spectral efficiency, we can set the channel spacing equal or close to the baud rate. An interesting approach based on this principle, known as Nyquist WDM, has appeared in [6], necessitating, however, the generation of steep signal spectrums. This requirement can be achieved either by Nyquist-like optical filtering or by high-speed digital-to-analog converters (DACs) that enable electrical shaping [7]. While substantial progress has been done in those fields [8,9], it could be desirable to use less aggressive

Manuscript received September 3, 2013; revised January 17, 2014; accepted January 17, 2014; published February 21, 2014 (Doc. ID 196820).

The authors are with Institut Mines-Telecom/Telecom SudParis, CNRS UMR 5157 SAMOVAR, 91011 Evry Cedex, France (e-mail: frederic.lehmann@it-sudparis.eu).

http://dx.doi.org/10.1364/JOCN.6.000315

optical filtering or less demanding DAC for cost-reduction purposes. Furthermore, a potential deviation of laser source emission frequency as well as filter center frequency may also yield additional linear cross talk between channels in a Nyquist WDM configuration. For example, the ITU-T G.692 standard recommendation [10] indicates a maximum frequency deviation of the optical source of a fifth of the channel spacing. Thus in the case of a data rate around 30 Gbaud, the deviation may reach 6 GHz and will be a potential source of linear cross talk between Nyquist WDM channels. Alternatively, we propose to use low-cost, commercially available transmit shaping filters with roll-off factors ≥ 0.1 and channel spacing down to the baud rate. Since our numerical results are obtained through simulations, we restrict ourselves to a digital implementation of the transmit filters. In an experimental setup, a comparison with optical spectral shaping could also be of interest. In this approach, the existence of cross talk between the WDM channels [11] exacerbates the challenge to achieve reliable communications. The reason is that ISI and cross talk between channels generates two-dimensional (2D) interference that must be handled effectively by DSP at the receiver. Digital compensation of those impairments is potentially attractive, since according to Moore's law the number of transistors per unit surface of silicon is doubling approximately every 2 years.

For 1D interference (i.e., ISI only), the benefit of allowing interaction between equalization and FEC decoding for optical transmissions is well documented [12]. However, this assumes that the effect of cross talk between WDM channels can be ignored, which is not the case in the scenario of interest in the present paper. To address this problem, some form of joint processing of the spectrally overlapping asynchronous WDM channels is required. In the recent past, a number of advanced techniques have been proposed in the literature [13–15]. The key feature common to these algorithms is their ability the exploit FEC decoding to refine 2D interference (i.e., both ISI and cross talk) cancellation iteratively. In the absence of ISI, [13] cancels the strong cross talk between channels using linear filtering. In [14], ISI mitigation is performed using the maximum *a posteriori* probability (MAP) algorithm [16], while cross talk is handled using parallel interference cancellation (PIC) [17]. The drawback of this approach is the exponential complexity increase in the modulation order and the size of the 1D ISI state. In [15], a 2D state-space representation is introduced to model ISI and cross talk in a unified way and cancellation is performed using Gaussian belief propagation (GaBP) [18]. An interesting property of this technique is that its complexity is independent of the modulation order. Note that 2D interference cancellation without interaction with FEC decoding, using either MAP detection [19,20] or adaptive filtering [20–23], has also appeared in the context of optical communications.

We consider methods assuming linear propagation effects with perfect knowledge of the channel parameters. Therefore, the issue of channel estimation must be investigated in order to achieve a practical implementation for the problem of interest. We derive the channel estimation algorithms relying on a maximum-likelihood (ML) approach. In order to keep the complexity at a reasonable level, carrier phase recovery is implemented in the form of a decision-directed (DD) phase-locked loop (PLL), which exploits delayed hard decisions taken after 2D interference cancellation. Also, a training-based method is introduced to estimate the 2D interference coefficients and noise variance for each WDM channel.

The main technical contributions of this work are

- the assessment of code-aided ISI and cross-talk (2D) cancellation, obtained through iterative message passing between FEC decoding and 2D interference cancellation.
- a phase estimation method working jointly with 2D interference cancellation able to resolve phase ambiguities inherent to quadrature amplitude modulation/phase-shift keying (QAM/PSK) modulations, thus avoiding the need for differential encoding.
- a 2D channel coefficient estimation procedure requiring only a limited number of pilot symbols per subchannel. This feature is required to complete the estimation of the constant channel parameters before the phase rotation induced by oscillator phase noise becomes harmful.
- spectrally efficient joint processing of multiple WDM subchannels not limited to quadrature phase-shift keying (QPSK) modulation.

This paper is organized as follows. First, Section II models the linear propagation effects as a noncausal 2D interference problem, which is in turn cast in a state-space form. In Section III, channel estimation algorithms based on the ML principle are derived. In Section IV, joint interference cancellation and decoding methods are discussed. Finally, in Section V, the performances of the proposed algorithms are assessed through numerical simulations on a long-haul fiber channel model taking into account linear and nonlinear effects.

II. Linear Fiber Model

A. System Setup

We consider the transmission over a long-haul distance of I WDM subchannels with channel spacing equal to the baud rate.

The transmitter (Tx) side is illustrated by Fig. 1. Time-frequency coding across the subchannels is used to improve the resilience to ISI and cross talk in such ultradense WDM systems. The binary information of all subchannels is first multiplexed, then encoded by a single high-rate FEC encoder [24,25], and subsequently demultiplexed toward the different subchannels. The encoded bitstream of the ith subchannel, where $0 \leq i \leq I-1$, is converted to complex symbols $d_{i,j}$, where the discrete time instant $j = 0, \ldots, J-1$, using a size-M QAM or PSK alphabet. In this paper, we assume that each subchannel is shaped with a commercially available root raised-cosine (RRC) filter with roll-off β.

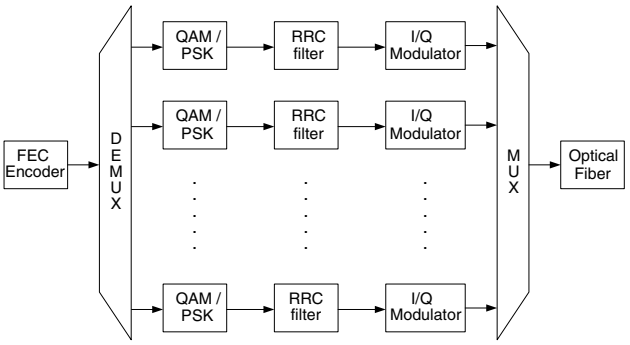

Fig. 1. Transmitter configuration for the coherent WDM system.

The receiver (Rx) side is illustrated by Fig. 2. After coherent demodulation, the signal of each subband is low-pass filtered in the electrical domain. Then, an analog-to-digital convertor (ADC) samples the signal of each subband with ideal timing recovery, followed by an electronic dispersion compensator (EDC) [1] implementing CD compensation in the frequency domain.

B. System Model

At the Tx side, the modulated symbols can be represented in matrix form, where $d_{i,j}$ denotes the complex symbol sent over the ith subband at the discrete time instant j. As illustrated in Fig. 3, the first N_t symbols on each subband correspond to training symbols known to the receiver. As a matter of fact, here we introduce a training sequence for each codeword only for the sake of averaging out the effect of asynchronism between WDM channels, as well as channel estimation errors during Monte Carlo simulations. It is understood that in real systems, the optical fiber parameters are very slowly time-varying, so that in practice training sequences are needed during startup and after that only once in a while.

We now model the noisy observations on the subchannels after EDC at the Rx side, $\{y_{i,j}\}$. The effect of the

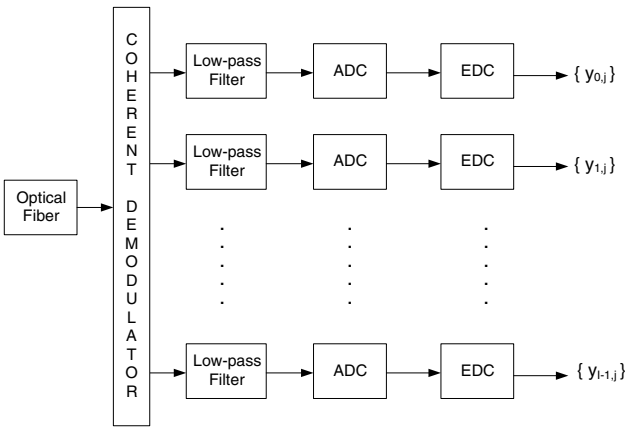

Fig. 2. Receiver layout for the coherent WDM system.

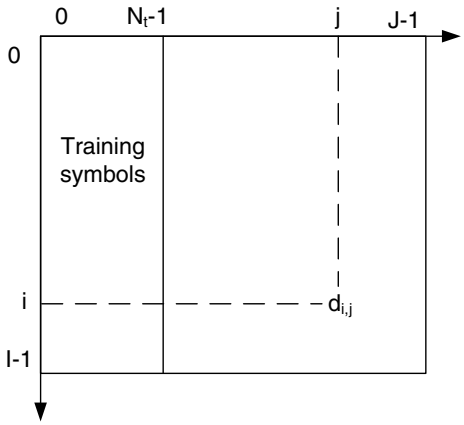

Fig. 3. Matrix of transmitted training and data symbols on each subband.

optical fiber include linear channel impairments (ISI due to imperfect EDC and cross talk) and the unknown carrier phase due to the receiver local oscillator (LO). Nonlinear fiber effects are left unmodeled but are taken into account in simulation results. Assuming that the Rx low-pass filter bandwidth B is sufficiently small, each subband experiences cross talk only from two adjacent subbands. Therefore, the EDC outputs can be modeled as the output of a 2D interference channel, and for $0 \le i \le I-1$, $0 \le j \le J-1$,

$$y_{i,j} = e^{j\theta_i} \sum_{m=-1}^{1} \sum_{n=-N/2}^{N/2} f_{m,n}^{i} d_{i-m,j-n} + n_{i,j}, \tag{1}$$

where $N+1$ is the ISI length, θ_i is the carrier phase, and $\{f_{m,n}^{i}\}$ is the set of complex 2D channel coefficients for the ith subband. These channel coefficients quantify the resulting interference between symbols of a subband or symbols on adjacent subbands stemming from linear optical propagation phenomena. The nonlinear impairments taken into account in the optical channel simulations will be detailed in Section V. $n_{i,j} \sim \mathcal{N}(0, \sigma_i^2)$ is an additive white Gaussian noise (AWGN) term, accounting for the presence of amplified spontaneous emission (ASE) background noise. Guard bands and guard intervals are inserted so that $d_{i-m,j-n} = 0$ in Eq. (1), if the condition $0 \le i-m \le I-1$ or $0 \le j-n \le J-1$ is violated. The complete system model is illustrated by Fig. 4.

Let us define the vector of conjugate 2D interference coefficients for the ith subchannel, of size $3(N+1) \times 1$, as

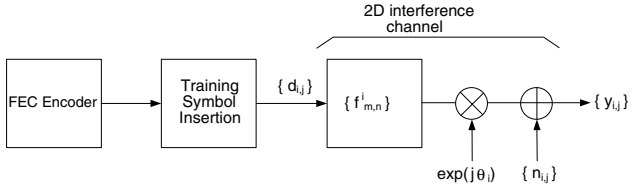

Fig. 4. WDM system model.

$$\mathbf{f}^i = \begin{bmatrix} {f^i_{1,N/2}}^* \\ {f^i_{0,N/2}}^* \\ {f^i_{-1,N/2}}^* \\ \vdots \\ {f^i_{1,-N/2}}^* \\ {f^i_{0,-N/2}}^* \\ {f^i_{-1,-N/2}}^* \end{bmatrix}. \tag{2}$$

The parameters of the 2D channel for the ith subchannel can be collected in $\lambda_i = (\mathbf{f}^i, \sigma_i^2)$. Also, define the vector of symbols affecting $y_{i,j}$ as

$$\mathbf{x}^i_j = \begin{bmatrix} d_{i-1,j-N/2} \\ d_{i,j-N/2} \\ d_{i+1,j-N/2} \\ \vdots \\ d_{i-1,j+N/2} \\ d_{i,j+N/2} \\ d_{i+1,j+N/2} \end{bmatrix}. \tag{3}$$

It follows that Eq. (1) can be rewritten as

$$y_{i,j} = e^{j\theta_i} \mathbf{f}^{iH} \mathbf{x}^i_j + n_{i,j}, \qquad 0 \le i \le I-1, 0 \le j \le J-1. \tag{4}$$

Remark 2.1: Due to the existence of guard bands, subband 0 (respectively, $I-1$) has no cross talk with a previous (respectively, a next) subband, therefore vectors (2) and (3) must be truncated accordingly.

III. 2D Channel Estimation

In this section, we derive ML estimation algorithms for the carrier phase and the 2D channel parameters in each subband.

A. Carrier Phase Estimation

The synthesis of the ML carrier phase estimation for the ith subband is based on the maximization of the log-likelihood function [26]. According to Eq. (4), for $0 \le j \le J-1$ we have

$$\ln p(y_{i,j}|\theta_i, \mathbf{x}^i_j, \lambda_i) = \ln \frac{1}{\pi\sigma_i^2} - \frac{1}{\sigma_i^2} |y_{i,j} - e^{j\theta_i} \mathbf{f}^{iH} \mathbf{x}^i_j|^2, \tag{5}$$

or equivalently

$$\ln p(y_{i,j}|\theta_i, \mathbf{x}^i_j, \lambda_i) = \ln \frac{1}{\pi\sigma_i^2} - \frac{1}{\sigma_i^2} (|y_{i,j}|^2 + |\mathbf{f}^{iH} \mathbf{x}^i_j|^2) + \frac{2}{\sigma_i^2} \mathrm{Re}((\mathbf{f}^{iH} \mathbf{x}^i_j)^* y_{i,j} e^{-j\theta_i}). \tag{6}$$

We can discard the two first terms in Eq. (6), since they are independent of the variable of interest, θ_i. Therefore, the ML estimation of θ_i reduces to the maximization of the objective function

$$L(y_{i,j}|\theta_i, \mathbf{x}^i_j, \lambda_i) = \mathrm{Re}((\mathbf{f}^{iH} \mathbf{x}^i_j)^* y_{i,j} e^{-j\theta_i}). \tag{7}$$

Since the objective function is a concave function of θ_i, the maximization can be performed using the stochastic gradient algorithm [27], assuming that the initial estimate is within the convergence region. Let $\hat{\theta}_{i,j}$ be the carrier phase estimate of the ith subband at the discrete time instant j, we obtain the update rule

$$\hat{\theta}_{i,j+1} = \hat{\theta}_{i,j} + \gamma_i \left. \frac{\partial L(y_{i,j}|\theta_i, \mathbf{x}^i_j, \lambda_i)}{\partial \theta_i} \right|_{\theta_i = \hat{\theta}_{i,j}},$$

where the step size γ_i must be optimized in order to track the carrier phase variations due to the Rx LO linewidth. Finally, using Eq. (7), the ML estimation algorithm has the form of a PLL:

$$\hat{\theta}_{i,j+1} = \hat{\theta}_{i,j} + \gamma_i \, \mathrm{Im}((\mathbf{f}^{iH} \mathbf{x}^i_j)^* y_{i,j} e^{-j\hat{\theta}_{i,j}}). \tag{8}$$

Since the channel parameters are not known *a priori* at the Rx side, $\mathbf{f}^i$ must be replaced by its estimate derived in Subsection III.B. Also, the data symbols contained in $\mathbf{x}^i_j$ are unknown to the receiver except during the training phase. Thus, we adopt a DD strategy, where hard symbol decisions, computed during the interference cancellation stage, are used instead of the true symbols. Note that a decision delay δ is necessary for two reasons: first because $\mathbf{x}^i_j$ is anticausal [see Eq. (3)] and second because in the retained interference cancellation algorithms (see Section IV), hard decisions become reliable only after a certain lag.

B. 2D Channel Parameter Estimation

We derive a ML estimation algorithm for the 2D channel parameters λ_i, based on a data-aided (DA) strategy during the initial training phase containing N_t symbols. The derivation is given for subband i, where $1 \le i \le I-2$. For subbands 0 and $I-1$, similar expressions are straightforwardly obtained based on remark 2.1.

Assuming N_t is small enough, so that the carrier phase θ_i remains nearly constant during the training phase, the term $e^{j\theta_i}$ can be absorbed into the complex channel vector $\mathbf{f}_i^H$ in Eq. (4). Thus, by setting the carrier phase estimate to zero at the end of the training phase, the PLL developed in Subsection III.A is automatically properly initialized.

Let T denote the subset of discrete time indices, where the noisy observations in each subband depend exclusively

on training symbols. Let $\mathbf{Y}_i = \{y_{i,j}, j \in T\}$ and $\mathbf{D}_i = \{\mathbf{x}_j^i, j \in T\}$. The desired DA ML estimate is given by

$$\hat{\lambda}_i = \arg\max_{\lambda_i} \log p(\mathbf{Y}_i|\mathbf{D}_i;\lambda_i), \tag{9}$$

which admits the following closed form expression:

$$\begin{cases} \hat{\mathbf{f}}^i & = \left[\sum_{j\in T} \mathbf{x}_j^i \mathbf{x}_j^{iH}\right]^{-1}\left[\sum_{j\in T} y_{i,j}^* \mathbf{x}_j^i\right] \\ \hat{\sigma}_i^2 & = \frac{\sum_{j\in T}|y_{i,j}-\hat{\mathbf{f}}^{iH}\mathbf{x}_j^i|^2}{|T|}. \end{cases} \tag{10}$$

The proof is postponed to Appendix A.

IV. Joint 2D Interference Cancellation and Decoding

Consider the EDC output at the Rx side, for the ith subchannel. We assume that the 2D channel parameters $\{f_{m,n}^i\}$ are constant during a codeword, so that the estimates obtained in Subsection III.B need not be updated after the initial training phase. Our optical fiber simulations have shown that this assumption holds true. Also, the PLL obtained in Subsection III.A compensates for the carrier phase drift due to the Rx LO by multiplying Eq. (1) with the inverse phase rotation $e^{-j\hat{\theta}_{i,j}}$.

The remaining task consists of reliable data detection on all WDM subchannels. In order to efficiently mitigate the effect of cross talk and residual ISI, we propose to apply 2D interference cancellation algorithms. Then, FEC decoding across the subchannels is used to correct the remaining errors. In this section, we present three different approaches to complete these tasks.

A. Blind MIMO Equalization

Blind multiple-input multiple-output (MIMO) equalization will be used as a benchmark method, since it is already widely used in coherent optical transmissions to compensate CD and PMD [1]. The most commonly used technique for blind MIMO equalization is the adaptive constant modulus algorithm (CMA) [28]. Since the CMA is a blind method, the channel estimation algorithms developed in Section III are not needed. However, a well-known drawback of adaptive CMA equalization is its slow convergence rate. In order to alleviate the slow convergence problem, we iterate the equalizer back and forth on each data block [29] until no further improvement is observed. Phase and delay ambiguities are other issues inherent to CMA that can be solved by using a known preamble and differential encoding at the Tx side [30]. The corresponding DSP processing at the Rx side is illustrated in Fig. 5. First, fractionally spaced blind MIMO equalization is applied to the EDC outputs. After noncoherent demodulation, bit-by-bit probabilities are calculated and fed to a SISO FEC decoder. In this setting, MIMO equalization is performed only once, since the CMA cannot refine its output by exploiting the FEC decoder's output.

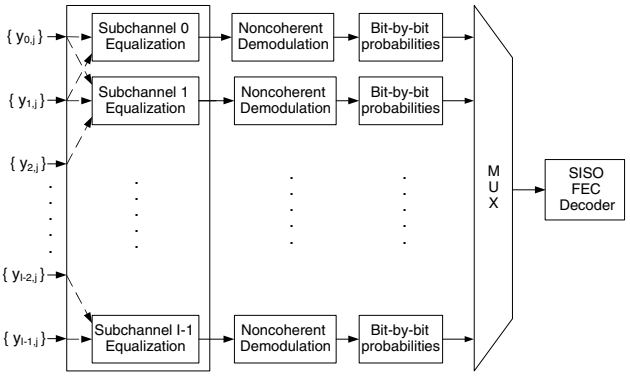

Fig. 5. WDM interference cancellation using blind MIMO equalization.

Let L denote the number of taps in each subchannel equalizer in Fig. 5. Then the complexity of blind MIMO equalization per complex data symbol is $\mathcal{O}(L)$. Note that the complexity is independent of the alphabet size M.

B. Per-Subchannel MAP/PIC Equalization

A state-of-the-art method for 2D interference cancellation was introduced in [14]. The corresponding iterative DSP processing at the Rx side is illustrated in Fig. 6. The cross talk from adjacent subbands is assimilated to a complex Gaussian random variable, whose mean and variance are estimated from the FEC decoder soft output and the 2D channel parameter estimates (see Subsection III.B). At each iteration, PIC [17] is applied to the EDC outputs, so that the remaining impairments are ISI, noise, and residual cross talk. Therefore, per-subband processing becomes feasible. First, the carrier phase is compensated with the PLL proposed in Subsection III.A. A slight modification is needed in Eq. (8), though. Indeed, the channel parameters in $\mathbf{f}^i$ corresponding to cross talk

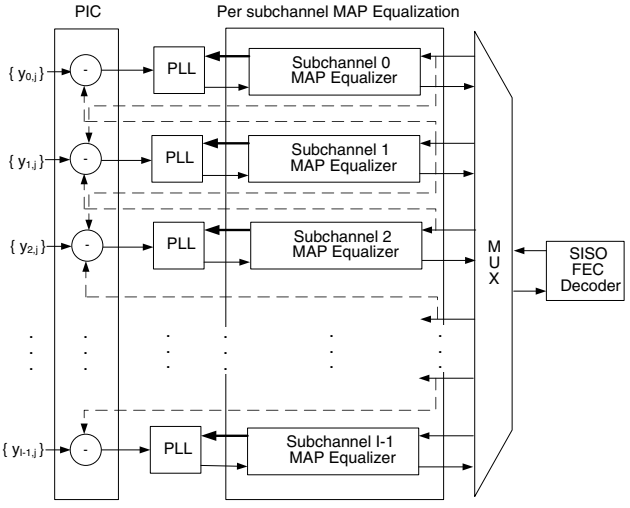

Fig. 6. WDM interference cancellation using per-subband MAP/PIC equalization with code-aided interference estimates (dashed lines) and PLL decision feedback loops (thick lines).

must be nulled, since the cross talk is assumed to be already cancelled by PIC. Then, each subband performs ISI mitigation using the MAP algorithm [16]. The conditional likelihood of the observations in the MAP algorithm is a Gaussian density, whose mean is based on both the ISI state and estimated channel parameters and whose variance is based on both the AWGN variance and estimated cross-talk variance. Since the MAP algorithm naturally generates soft decisions, they are fed to a SISO FEC decoder for the sake of error correction. This process is iterated until convergence is reached. At the initialization, the cross-talk mean and variance affecting each subband are computed assuming equiprobable data symbols in adjacent subbands.

The complexity of MAP/PIC equalization per complex data symbol and per iteration is $\mathcal{O}(M^N)$ [16]. Therefore, this method is tractable only for a small alphabet size M and a small ISI length.

C. GaBP-Based 2D Interference Cancellation

In contrast to the two previous methods borrowed from the equalization literature, we now discuss a SISO algorithm tailored for 2D interference cancellation. The corresponding iterative DSP processing at the Rx side is illustrated in Fig. 7. First, the carrier phase is compensated with the PLL proposed in Subsection III.A. Then 2D interference cancellation is performed with the GaBP-based method introduced in [15], using the channel parameter estimation procedure in Subsection III.B.

GaBP-based 2D interference cancellation is an iterative SISO detector, which takes advantage of the fact that each row of observations corresponding to a given subband depends only on a strip of symbols transmitted by the current and two adjacent subbands. During one iteration, each row

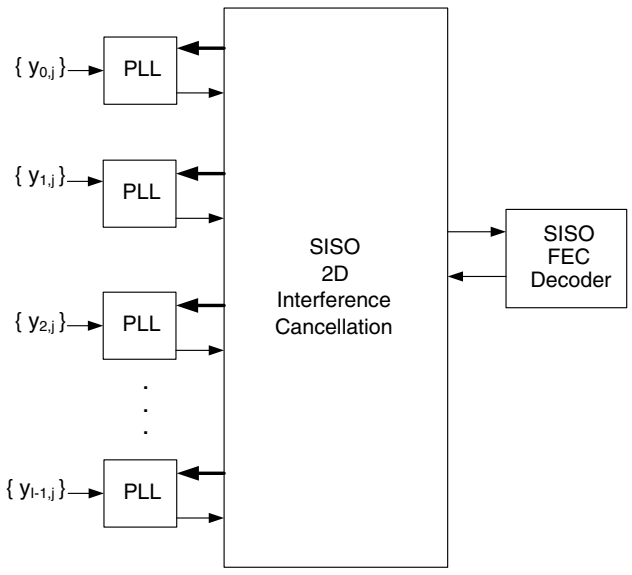

Fig. 7. WDM 2D interference cancellation using GaBP with PLL decision feedback loops (thick lines).

of observations is independently processed with a Kalman smoother equalizer [15], generating soft decisions for all the rows of symbols in the corresponding strip at a time. Since the strips corresponding to neighboring subbands overlap, multiple soft decisions are generated for each symbol. The equality constraints between the symbols in overlapping strips are then exploited in a message-passing fashion [18,31] during the next iteration. Since the algorithm in [15] is naturally SISO, soft decisions can be exchanged iteratively with a SISO FEC decoder to improve the reliability of the decisions until convergence is reached.

The complexity of GaBP-based interference cancellation was evaluated in [15]. We consider only cross talk from two adjacent subbands, so the complexity per complex data symbol and per iteration becomes $\mathcal{O}((3(N+1))^3)$, which is independent of M. Therefore, GaBP-based interference cancellation is applicable for any alphabet size M.

V. Simulation Results

At the Tx side, each codeword is generated using a length-3000, regular LDPC code [31] with coding rate 0.8. A codeword is demultiplexed to $I = 5$ WDM subbands with spacing equal to the baud rate $1/T$, where T denotes the symbol duration, also using an oversampling factor of 64. Regarding the size of the modulation alphabet, $M = 4$, 8, or 16. $N_t = 60$ training symbols are inserted for each subband. RRC shaping with roll-off $\beta = 0.1$ is selected in order to demonstrate the potential of the proposed 2D interference cancellation methods while keeping the cross talk at a reasonable level. The guard band between two consecutive WDM multiplexes is set to β/T Hz in order to eliminate any interference. Similarly, a guard interval of five symbols is inserted between two consecutive frames in each subband.

At the Rx side, we consider state-of-the-art LOs with laser linewidth of 100 kHz [32]. All low-pass filters are order-5 Bessel filters, with bandwidth set to $B = 0.7/T$ in order to limit the cross talk to two adjacent subbands. The delay in the DD PLLs is fixed to $\delta = N + N/2$ in order to feed back reliable hard decisions, and the step size is optimized to 7×10^{-2}. SISO FEC decoding is performed by applying the sum-product decoder [31]. Each time interference cancellation is completed, the decoder is activated and performs N_i inner iterations. The number of outer iterations, N_o, denotes the number interference cancellation iterations.

Regarding the blind MIMO equalization method, a $T/4$ fractionally spaced CMA algorithm is used [28]. Also, the training symbols are exploited only to resolve delay ambiguities. Moreover, as blind MIMO equalization is not a SISO method, N_o must be set to 1. T-spaced sampling is used for the per-subchannel MAP/PIC and the GaBP methods.

The performances are evaluated in terms of average bit-error rate (BER) after decoding as a function of the electrical signal-to-noise ratio (SNR).

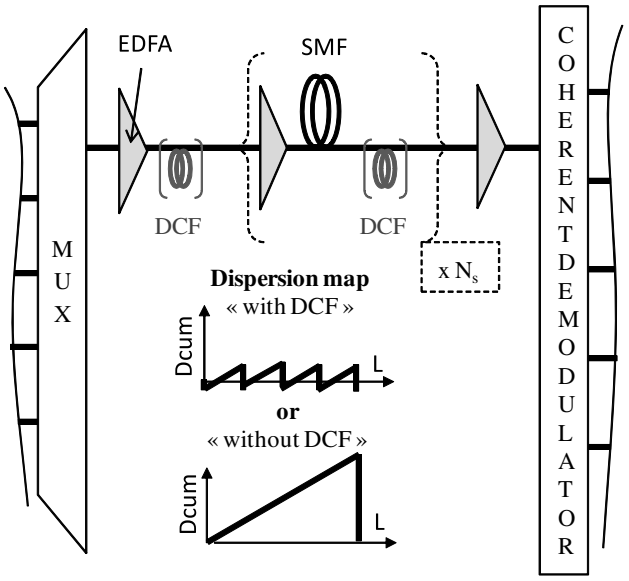

Fig. 8. WDM transmission line setup.

The transmission line setup used in numerical simulations is shown in Fig. 8. It consists of a succession of 100 km long standard single-mode fiber (SMF) spans with the number of spans N_s equal to 20. The fiber propagation is modeled using the split-step Fourier method taking into account only the loss, the group velocity dispersion (GVD), and the Kerr nonlinearities while dispersion slope, PMD, and Raman or Brillouin nonlinear effects are neglected here. The SMF parameters are thus an attenuation coefficient $\alpha = 0.2$ dB/km, a GVD parameter $D = 17$ ps $\cdot$ nm^{-1} $\cdot$ km^{-1}, a nonlinear index $n_2 = 2.7 \cdot 10^{-20}$ m^2/W, and an effective area $A_{\text{eff}} = 80$ μm^2. We consider an optical amplification between fiber spans made by idealized flat-gain and noiseless erbium-doped fiber amplifiers (EDFAs) yielding the same launched average signal power per channel (referred to as $P_{\text{in,avg/ch}}$) at the beginning of each fiber span. The ASE noise of EDFA is taken into account by applying an equivalent noise loading for a targeted optical SNR at the receiver side before counting the errors using the Monte Carlo method. We thus neglect the nonlinear signal-to-noise interaction (NSNI) that may occur during the nonlinear propagation. Such NSNI has been shown to reduce the nonlinear threshold of a transmission by about 1 dB in [33] for our considered baud rate of 32 Gbaud with channel frequency spacing of 32 GHz. In order to analyze the efficiency of our DSP scheme at the receiver side, we perform the transmission line simulations in two steps, first assuming linear propagation (no nonlinear Kerr index in the modeled SMF) in Subsection V.A and second also applying the Kerr nonlinearity in the fiber model in Subsection V.B. For the linear fiber model, we choose to compensate for the cumulative GVD during the fiber propagation by the unique use of EDC within the DSP blocks, while for the nonlinear case we consider two dispersion management (DM) configurations. To illustrate these last two options, Fig. 8 presents the two corresponding dispersion maps yielding the cumulative dispersion D_{cum} as a function of the transmission distance L. One scheme referred to as the DM configuration without dispersion-compensating fiber (DCF) does not incorporate in-line dispersion compensation between fiber spans. The GVD compensation is then performed electronically before or after the transmission line. This scheme has been shown to yield the best performances for coherent transmission systems of signal with amplitude and phase modulated at 32 Gbaud. The second scheme referred to as the DM configuration with DCF includes in-line DCF that fully compensates for the cumulative dispersion of the SMF previous span. In our simulation we assume the $P_{\text{in,avg/ch}}$ is sufficiently low to neglect the nonlinearity in the DCF. We may also consider (if this is specified) a precompensation of GVD applied using an additional linear DCF after the transmitter. This last DM scheme was revealed to be the best case to mitigate nonlinear effects for 10 Gbit/s intensity modulation and direct detection (IM/DD) legacy systems.

A. Linear Optical Fiber

In this subsection, we consider as mentioned previously a linear transmission using SMF without in-line dispersion compensation. By choosing this setup, we ensure that the system model introduced in Subsection II.B is correct. We perform Monte Carlo simulations by encoding random information bits for each frame. For each new codeword, the delay between the subchannels accounting for asynchronism is randomly drawn in an interval of size equal to the symbol duration. Therefore, the 2D channel parameters need to be re-estimated for each codeword. This is clearly a conservative approach, since we have verified that the channel parameters on a standard optical fiber are constant for a longer time, but this method has the advantage of averaging out the effect of asynchronism between subbands.

Due to imperfect EDC, the length of the residual ISI on each subchannel must be optimized: it must not be too high in order to avoid overfitting during parameter estimation, and it cannot be too small, otherwise an error floor due to unmodeled ISI appears. Here, we found that N should be set to 4 for the MAP/PIC and GaBP-based methods, while the MIMO equalizer needs a 81×3-tap filter per subchannel.

Figure 9 compares the BER of blind MIMO equalization and iterative GaBP interference cancellation for the same number of FEC decoder iterations. Note that the benefit of MIMO-CMA over per-subcarrier CMA equalization has been demonstrated previously in [21]. It is common practice to seek BERs around 10^{-3} before FEC correction and BERs around 10^{-12} after FEC correction. For the sake of reasonable time consumption, here we perform simulations only down to BERs around 10^{-5} to get the beginning of the evolution of each BER curve as a function of the SNR using FEC, since no error floor is encountered at higher SNR for long regular LDPC codes [34]. Thus we consider it relevant to predict the BER evolution toward higher SNR values by extrapolation. No BER improvement could be obtained by further increasing the values of N_o or N_i.

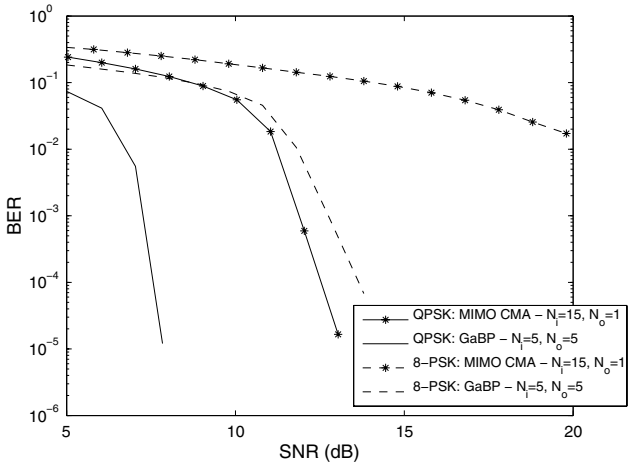

Fig. 9. BER comparison of blind MIMO equalization and GaBP 2D interference cancellation.

Blind MIMO equalization suffers from an SNR loss due to the suboptimality of noncoherent demodulation and also from residual cross talk and ISI. The good performances of GaBP interference cancellation are mainly due to iterative soft decision feedback from the FEC decoder. Figure 10 compares the BER of per-subchannel MAP/PIC equalization and iterative GaBP interference cancellation for the same number of inner (FEC) and outer iterations. The performances are almost identical for QPSK modulation, while per-subchannel MAP/PIC equalization has a 1 dB advantage over iterative GaBP interference cancellation for 8-QAM modulation at a BER of 10^{-4}. However, the computational complexity of per-subchannel MAP/PIC equalization is very high for 8-QAM modulation, since each constituent MAP equalizer works on a trellis with $8^4 = 4096$ states. For 16-QAM modulation, it was not even possible to run per-subchannel MAP/PIC equalization with the available computational resources (due to the use of trellises with $16^4 = 65536$ states), while there was no problem for iterative GaBP interference cancellation. This confirms

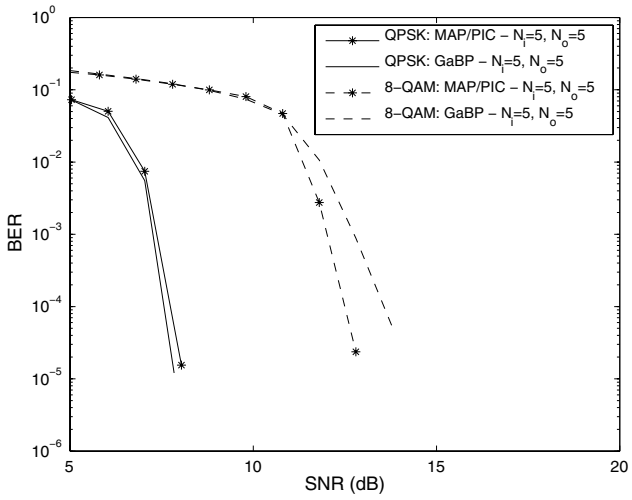

Fig. 10. BER comparison of per-subchannel MAP/PIC equalization and GaBP 2D interference cancellation.

the exponential complexity increase of MAP/PIC and the relative insensitivity of the GaBP-based method with respect to the modulation order.

Another interesting issue is how much can be gained from the type of advanced DSP processing at the Rx side that we advocate. Figure 11 compares the proposed iterative GaBP 2D interference cancellation with a per-subchannel GaBP processor taking care of ISI while ignoring the existence of WDM cross talk. The proposed GaBP-based 2D interference cancellation method is 0 and 2.5 dB more power efficient than per-subchannel GaBP processing at a BER of 10^{-4} for QPSK and 8-QAM, respectively. For 16-QAM, per-subchannel GaBP reaches a high error floor. These results can be interpreted as follows. During parameter estimation, per-subchannel GaBP processing considers only the ISI coefficients. Therefore, when computing the noise variance estimate, WDM cross talk is simply treated as additional Gaussian noise. This coarse approximation is fine for QPSK, grossly suboptimal for 8-QAM, and eventually breaks down completely for higher order modulations.

We also provide the BER of per-subchannel GaBP processing taking care of ISI, with genie-aided ideal cross-talk removal, which acts as a lower bound on the achievable performances. Note that this lower bound could only be reached for Nyquist WDM at the expense of ideal filters with rectangular spectrum. We observe that the proposed GaBP 2D interference cancellation method is only 1, 1.5, and 2 dB away from that lower bound at a BER of 10^{-4} for QPSK, 8-QAM, and 16-QAM, respectively.

Before comparing DSP techniques for 2D interference cancellation in the nonlinear regime, we also want to indicate the potential influence of frequency offsets among WDM subbands. For instance, [21] indicates that MIMO/CMA-based DSP seems robust toward potential LO frequency offsets, while [23] shows that least-mean-square-based DSP remains sensitive to those offsets. For the proposed algorithms, this aspect is outside the

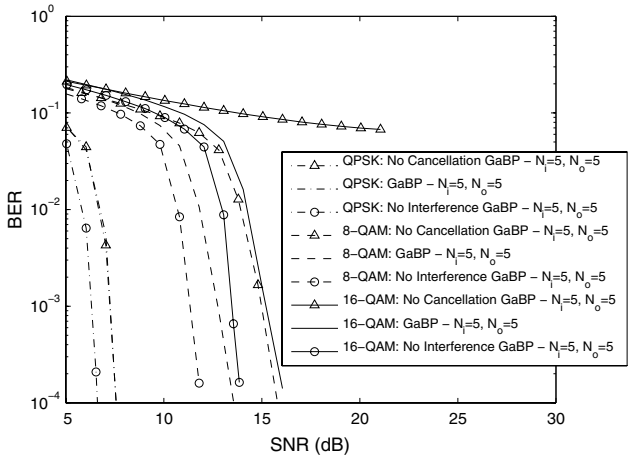

Fig. 11. BER comparison of per-subchannel GaBP ignoring the presence of WDM cross talk (triangles) and GaBP 2D interference cancellation (lines). The BER of a genie-aided per-subchannel GaBP with artificial cross-talk removal is also provided (circles).

scope of the present work and is left as a subject of future investigation.

B. Nonlinear Optical Fiber

The proposed Rx DSP algorithms are based on a linear optical fiber model. Therefore, it is necessary to check the sensitivity of the proposed algorithms to nonlinear effects. We consider the nonlinear Kerr effect in the propagation model of the fiber and present the results for $P_{\text{in,avg/channel}} = -7$ dBm and 8-QAM modulation for $N_s = 20$ spans. Here, we implement ideal EDC and found that the best ISI length parameter should be set to $N = 2$, both for MAP/PIC and GaBP 2D interference cancellation.

BER results are given in Fig. 12 for the DM configuration with DCF. The BER curve of per-subchannel GaBP processing taking care of ISI, with genie-aided ideal cross-talk removal, constitutes again a lower bound on the achievable performances. Although not shown here, the BER curve of per-subchannel MAP processing taking care of ISI, with genie-aided ideal cross-talk removal, is almost identical. At a BER of 10^{-4}, the MAP/PIC and the GaBP 2D interference cancellation are, respectively, 3 and 5 dB away from the lower bound, due to the presence of WDM nonlinear cross talk. Comparing with the results obtained for 8-QAM on a linear optical fiber in Fig. 10, we note that the MAP/PIC method is more robust against fiber nonlinearities than GaBP 2D interference cancellation. This added robustness comes at a price of a higher computational complexity, though.

Figure 13 shows similar results for the DM configuration without DCF. At a BER of 10^{-4}, the MAP/PIC and the GaBP 2D interference cancellation are, respectively, 2 and 4 dB away from the lower bound, due to the presence of WDM nonlinear cross talk.

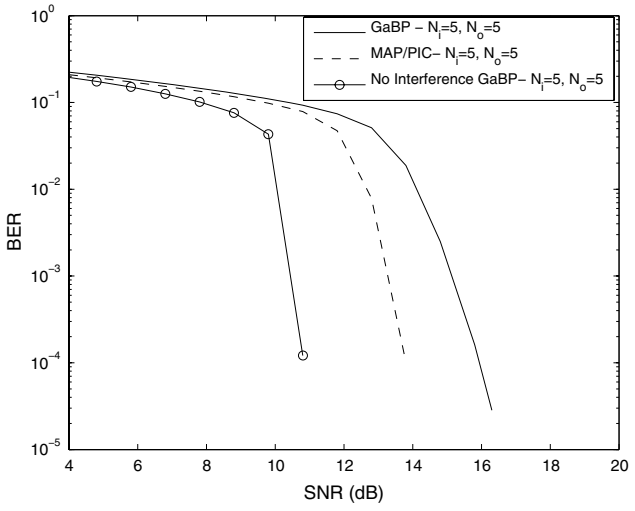

Fig. 12. BER comparison on a nonlinear fiber with DCF for 8-QAM: GaBP 2D interference cancellation (lines), per-subchannel MAP/PIC equalization (dashed), and genie-aided per-subchannel GaBP with artificial cross-talk removal (circles).

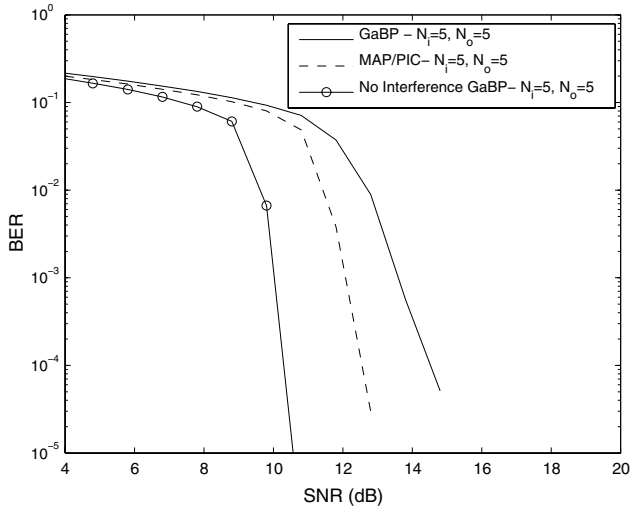

Fig. 13. BER comparison on a nonlinear fiber without DCF for 8-QAM: GaBP 2D interference cancellation (lines), per-subchannel MAP/PIC equalization (dashed), and genie-aided per-subchannel GaBP with artificial cross-talk removal (circles).

It follows that the DM configuration without DCF is 1 dB more power efficient than the DM configuration with DCF in the presence of nonlinear cross talk. This performance improvement is due to a reduced impact of fiber nonlinearities between channels during the propagation when considering system design without in-line compensation of CD. For example, it has been shown in [35] that the nonlinear threshold is 2 dB higher when using a system design without in-line DCF. If we further increase the input power $P_{\text{in,avg/channel}}$, the fiber nonlinearities become the dominant impairment, and the proposed algorithms exhibit an error floor. The analysis in the presence of fiber nonlinearity indicates that the processing algorithm under investigation here is not able to mitigate high impairments due to nonlinear ISI but remains efficient for the compensation of linear cross talk when the signals are launched with an input power close to the nonlinear threshold.

VI. Conclusions

In this paper, we have designed DSP methods to mitigate the effect of ISI and cross talk in WDM coherent optical transmissions. We have shown that 2D interference cancellation combined with FEC decoding at the receiver side can recover the information sent by a WDM multiplex with channel spacing down to the baud rate. Moreover, we have derived efficient estimation algorithms based on the ML principle for the purpose of channel estimation in nonblind methods. We have found that per-subband MAP/PIC and GaBP-based interference cancellation outperform blind MIMO equalization, which is currently the reference method for interference mitigation in optical communications. The key to effectiveness is the use of iterative soft decision feedback from a FEC decoder. We have also shown that a simple DD PLL, before interference cancellation, is sufficient for carrier phase recovery if we consider state-of-the-art local oscillators.

Future developments will include an extension to other forms of cross talk in optical fibers due to PolMux. If PolMux is used on top of WDM, the number of channels will be doubled in our system model, and each channel will interfere with more than two adjacent channels. A straightforward extension of the proposed algorithms is able to handle jointly ISI and the cross talk between polarization modes and neighboring subbands. Another interesting future direction would be to adapt the proposed scheme to the more involved problem of modal cross talk in coherent mode division multiplexed transmissions [36].

Appendix: Proof of Equation (10)

Due to the AWGN assumption in the system model of Subsection II.B, the conditional likelihood admits the following factorization:

$$p(\mathbf{Y}_i|\mathbf{D}_i;\boldsymbol{\lambda}_i) = \prod_{j\in T} p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i). \tag{A1}$$

Using the product rule, the gradient of Eq. (A1) with respect to $\boldsymbol{\lambda}_i$ is expressed as

$$\nabla_{\boldsymbol{\lambda}_i} p(\mathbf{Y}_i|\mathbf{D}_i;\boldsymbol{\lambda}_i) = p(\mathbf{Y}_i|\mathbf{D}_i;\boldsymbol{\lambda}_i) \sum_{j\in T} \frac{\nabla_{\boldsymbol{\lambda}_i} p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i)}{p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i)}, \tag{A2}$$

or equivalently

$$\nabla_{\boldsymbol{\lambda}_i} \log p(\mathbf{Y}_i|\mathbf{D}_i;\boldsymbol{\lambda}_i) = \sum_{j\in T} \nabla_{\boldsymbol{\lambda}_i} \log p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i). \tag{A3}$$

Thus, the solution of Eq. (9) verifies

$$\sum_{j\in T} \nabla_{\boldsymbol{\lambda}_i} \log p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i)\big|_{\boldsymbol{\lambda}_i=\hat{\boldsymbol{\lambda}}_i} = \mathbf{0}. \tag{A4}$$

Since the phase rotation $e^{j\theta_i}$ is absorbed into the complex channel vector $\mathbf{f}_i^H$ during the initial training phase, we have

$$p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i) = \frac{1}{\pi\sigma_i^2} \exp\left\{-\frac{1}{\sigma_i^2}|y_{i,j} - \mathbf{f}^{iH}\mathbf{x}_j^i|^2\right\}. \tag{A5}$$

Taking the gradient of the logarithm of Eq. (A5), with respect to the complex valued parameter vector $\mathbf{f}^i$ (see [27], p. 798), we obtain

$$\begin{aligned}\nabla_{\mathbf{f}^i} \log p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i) &= -\frac{1}{\sigma_i^2}\nabla_{\mathbf{f}^i}|y_{i,j} - \mathbf{f}^{iH}\mathbf{x}_j^i|^2 \\ &= -\frac{2}{\sigma_i^2}(\mathbf{x}_j^i\mathbf{x}_j^{iH}\hat{\mathbf{f}}^i - y_{i,j}^*\mathbf{x}_j^i),\end{aligned} \tag{A6}$$

where the second line is obtained by exploiting the expressions on p. 796 in [27]. Moreover, taking the partial derivative of the logarithm of Eq. (A5), with respect to the real coefficient σ_i^{-2}, we get

$$\frac{\partial \log p(y_{i,j}|\mathbf{x}_j^i;\boldsymbol{\lambda}_i)}{\partial\sigma_i^{-2}} = \sigma_i^2 - |y_{i,j} - \mathbf{f}^{iH}\mathbf{x}_j^i|^2. \tag{A7}$$

Injecting Eqs. (A6) and (A7) into Eq. (A4) completes the proof.

Acknowledgments

The authors would like to thank Amir Ghazisaeidi, with Alcatel-Lucent, France, for helpful discussions on the implementation of state-of-the-art Nyquist-like transmit filters.

References

[1] S. J. Savory, "Digital optical coherent receivers: Algorithms and subsystems," *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Sept./Oct. 2010.

[2] E. Ip, P. Ji, E. Mateo, Y.-K. Huang, L. Xu, D. Qian, N. Bai, and T. Wang, "100G and beyond transmission technologies for evolving optical networks and relevant physical-layer issues," *Proc. IEEE*, vol. 100, no. 5, pp. 1065–1078, May 2012.

[3] E. Yamazaki, M. Tomizawa, and Y. Miyamoto, "100-Gb/s optical transport network and beyond employing digital signal processing," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s43–s49, Feb. 2012.

[4] R. Le Bidan, C. Leroux, C. Jego, P. Adde, and R. Pyndiah, "Reed–Solomon turbo product codes for optical communications: From code optimization to decoder design," *EURASIP J. Wireless Commun. Netw.*, vol. 2008, 658042, 2008.

[5] K. Gracie and M. H. Hamon, "Turbo and turbo-like codes: Principles and applications in telecommunications," *Proc. IEEE*, vol. 95, no. 6, pp. 1228–1254, June 2007.

[6] G. Bosco, A. Carena, V. Curri, P. Poggiolini, and F. Forghieri, "Performance limits of Nyquist-WDM and CO-OFDM in high speed PM-QPSK systems," *IEEE Photon. Technol. Lett.*, vol. 22, no. 15, pp. 1129–1131, Aug. 2010.

[7] G. Bosco, "Spectral shaping in ultra-dense WDM systems: Optical vs. electrical approaches," in *Optical Fiber Communication Conf. (OFC)*, Los Angeles, CA, Mar. 2012.

[8] Y. Cai, J. X. Cai, C. R. Davidson, D. Foursa, A. Lucero, O. Sinkin, A. Pilipetskii, G. Mohs, and N. S. Bergano, "Achieving high spectral efficiency in long-haul transmission with prefiltering and multi-symbol detection," in *Proc. Asia Communications and Photonics Conf. and Exhibition (ACP)*, Shanghai, China, Dec. 2010, pp. 349–350.

[9] R. Schmogrow, M. Winter, M. Meyer, D. Hillerkuss, S. Wolf, B. Baeuerle, A. Ludwig, B. Nebendahl, S. Ben-Ezra, J. Meyer, M. Dreschmann, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold, "Real-time Nyquist pulse generation beyond 100 Gbit/s and its relation to OFDM," *Opt. Express*, vol. 20, no. 1, pp. 317–337, Jan. 2012.

[10] "Optical interfaces for multichannel systems with optical amplifiers," ITU-T Recommendation G.692, 1998.

[11] I. Tafur Monroy and E. Tangdiongga, *Crosstalk in WDM Communication Networks* (Kluwer Academic, 2002).

[12] I. B. Djordjevic, M. Arabaci, and L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.

[13] B. F. Beidas, H. El Gamal, and S. Kay, "Iterative interference cancellation for high spectral efficiency satellite

communications," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 31–36, Jan. 2002.

[14] N. Noels and M. Moeneclaey, "Spectrally efficient CPM: Suboptimal FG-based multiuser detection," in *Proc. Wireless Communications and Networking Conf. (WCNC)*, Paris, France, April 2012, pp. 918–923.

[15] F. Lehmann, "Iterative mitigation of intercell interference in cellular networks based on Gaussian belief propagation," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2544–2558, July 2012.

[16] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, pp. 284–287, Mar. 1974.

[17] S. Verdu, *Multiuser Detection*. Cambridge University, 1998.

[18] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[19] R. Thiruneelakandan, P. Watts, R. I. Killey, and M. Glick, "DSP-based crosstalk cancellation in WDM interconnects," in *Proc. CLEO/QELS*, San Jose, CA, May 2010.

[20] J. Pan, C. Liu, T. Detwiler, A. J. Stark, Y.-T. Hsueh, and S. E. Ralph, "Inter-channel crosstalk cancellation for Nyquist-WDM superchannel applications," *J. Lightwave Technol.*, vol. 30, no. 24, pp. 3993–3999, Dec. 2012.

[21] T. Zeng, "Superchannel transmission system based on multi-channel equalization," *Opt. Express*, vol. 21, no. 12, pp. 14799–14807, June 2013.

[22] C. Liu, J. Pan, T. Detwiler, A. Stark, Y.-T. Hsueh, G.-K. Chang, and S. E. Ralph, "Joint digital signal processing for superchannel coherent optical systems: Joint CD compensation for joint ICI cancellation," in *Proc. European Conf. and Exhibition on Optical Communications (ECOC)*, Amsterdam, The Netherlands, Sept. 2012.

[23] J. Pan, A. Stark, C. Liu, and S. E. Ralph, "Fractionally-spaced frequency domain linear crosstalk cancellation with spectral alignment techniques for coherent superchannel optical systems," in *Optical Fiber Communication Conf. and Exhibition and the Nat. Fiber Optic Engineers Conf. (OFC/NFOEC)*, Anaheim, CA, Mar. 2013.

[24] M. Mazurczyk, D. G. Foursa, H. G. Batshon, H. Zhang, C. R. Davidson, J.-X. Cai, A. Pilipetskii, G. Mohs, and N. S. Bergano, "30 Tb/s transmission over 6,630 km using 16QAM signals at 6.1 bits/s/Hz spectral efficiency," in *Proc. European Conf. and Exhibition on Optical Communications (ECOC)*, Amsterdam, The Netherlands, June 2012.

[25] D. Huang, K. Ben Letaief, and J. Lu, "Bit-interleaved time-frequency coded modulation for OFDM systems over time-varying channels," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1191–1199, July 2005.

[26] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*. Wiley, 1998.

[27] S. Haykin, *Adaptive Filter Theory.* Upper Saddle River, NJ: Prentice Hall, 2002.

[28] C. B. Papadias and A. J. Paulraj, "A constant modulus algorithm for multiuser signal separation in presence of delay spread using antenna arrays," *IEEE Signal Process. Lett.*, vol. 4, no. 6, pp. 178–181, June 1997.

[29] B.-J. Kim and D. C. Cox, "Blind equalization for short burst wireless communications," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1235–1247, July 2000.

[30] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[31] T. J. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.

[32] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Express*, vol. 16, no. 2, pp. 753–791, Jan. 2008.

[33] A. Bononi, P. Serena, and N. Rossi, "Nonlinear signal–noise interactions in dispersion-managed links with various modulation formats," *Opt. Fiber Technol.*, vol. 16, no. 2, pp. 73–85, Mar. 2010.

[34] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT, 1963.

[35] A. Bononi, P. Serena, N. Rossi, and D. Sperti, "Which is the dominant nonlinearity in long-haul PDM-QPSK coherent transmissions?" in *Proc. European Conf. and Exhibition on Optical Communications (ECOC)*, Torino, Italy, Sept. 2010.

[36] S. Randel, R. Ryf, A. Sierra, P. J. Winzer, A. H. Gnauk, C. A. Bolle, R.-J. Essiambre, D. W. Peckham, A. McCurdy, and R. Lingle, Jr., "6×56-Gb/s mode-division multiplexed transmission over 33-km few-mode fiber enabled by 66 MIMO equalization," *Opt. Express*, vol. 19, no. 17, pp. 16697–16707, Aug. 2011.

Frederic Lehmann received the E.E. and the M.S.E.E. degrees from ENSERG, France, in 1998. In 2002, he received the Ph.D. in electrical engineering from the National Polytechnical Institute, Grenoble (INPG), France. He worked as a research engineer with STMicroelectronics from 1999 to 2002. From 2003 to 2004, he was a postdoctoral researcher at the Laboratory for Analysis and Architecture of Systems (LAAS), CNRS, Toulouse, France. Currently, he is an assistant professor at the Institut TELECOM, Telecom SudParis, Evry, France. His main research interests are in the areas of communication theory, nonlinear signal processing, and statistical image processing.

Petros Ramantanis was born in Kavala, Greece, in 1982. He received a degree in electrical engineering from the National Technical University of Athens, Greece, in 2006; a masters degree in optics and photonics from Ecole Polytechnique, Palaiseau, France, in 2007; and a Ph.D. degree from Telecom SudParis, Evry, France, in 2011, for work on the performance prediction of coherent, dispersion-managed optical networks with advanced modulation. He is currently with the Optical Transmissions Group of Alcatel-Lucent Bell-Labs, Nozay, France.

Yann Frignac was born in 1975. He obtained his Ph.D. in electronics and communication in 2003 at Telecom ParisTech and Alcatel Research and Innovation on optical transmission systems. As a postdoctoral fellow at Commissariat l'Energie Atomique (2004–2005), he was in charge of research on fiber biophotonic sensors. Since 2006, he has been an associate professor at Telecom SudParis and CNRS Samovar (UMR5157), where he is in charge of optical transmission coherent system research including multi-level modulation formats, optical amplifiers, and multimode transmission modeling. He is the author or coauthor of 40 conference and peer-reviewed journal papers.