(article starts on next page)

Optics EXPRESS

Research Article

Overhead-optimization of pilot-based digital signal processing for flexible high spectral efficiency transmission

MIKAEL MAZUR,[1,*] JOCHEN SCHRÖDER,[1] ABEL LORENCES-RIESGO,[2] TSUYOSHI YOSHIDA,[3] MAGNUS KARLSSON,[1] AND PETER A. ANDREKSON[1]

[1]*Photonics Laboratory, Fibre Optic Communication Research Centre (FORCE), Department of Microtechnology and Nanoscience, Chalmers University of Technology, Gothenburg SE-412 96, Sweden*
[2]*IT-Instituto de Telecomunicações, 3810-193 Aveiro, Portugal*
[3]*Mitsubishi Electric Corp., Kamakura, Kanagawa, 247-8501, Japan*
*mikael.mazur@chalmers.se

Abstract: We present a low-complexity fully pilot-based digital signal processing (DSP) chain designed for high spectral efficiency optical transmission systems. We study the performance of the individual pilot algorithms in simulations before demonstrating transmission of a 51×24 Gbaud PM-64QAM superchannel over distances reaching 1000 km. We present an overhead optimization technique using the system achievable information rate to find the optimal balance between increased performance and throughput reduction from adding additional DSP pilots. Using the optimal overhead of 2.4%, we report 9.3 (8.3) bits/s/Hz spectral efficiency, or equivalently 11.9 (10.6) Tb/s superchannel throughput, after 480 (960) km of transmission over 80 km spans with EDFA-only amplification. Moreover, we show that the optimum overhead depends only weakly on transmission distance, concluding that back-to-back optimization is sufficient for all studied distances. Our results show that pilot-based DSP combined with overhead optimization can increase the robustness and performance of systems using advanced modulation formats while still maintaining state-of-the-art spectral efficiency and multi-Tb/s throughput.

1. Introduction

Today, optical networks use high symbol rate transceivers with advanced modulation formats to maximize the channel spectral efficiency (SE) [1]. These transceivers use all four available dimensions in single-mode fibers (SMFs) to transmit independent information, i.e. both I and Q on two orthogonal polarizations. To enable reliable communication and ensure a final bit error rate (BER) $< 10^{-15}$, advanced digital signal processing (DSP) [2] and forward error correction (FEC) [3] are needed. System impairments limiting the performance consist of both transceiver imperfections and distortions induced by the fiber channel. Transceiver imperfections include limited effective number of bits (ENOBs) [4, 5], limited bandwidth of both digital-to-analog converters (DACs) and analog-to-digital converters (ADCs) as well as nonlinear distortions [6, 7]. Distortions from transmission include amplified spontaneous emission noise from erbium-doped fiber amplifiers (EDFAs) and non-linear signal degradation from the fiber itself [8, 9].

As the throughput demands continue to grow, so does the requirement for flexibility; optical networks thus change from fixed point-to-point to optically routed flex-grid networks [1, 10]. Using techniques such as adaptive-rate FEC [11] and constellation shaping [12], individual transceivers can adapt the effective rate to maximize the performance over distances ranging from single- or few spans-links to submarine links connecting continents [13]. Depending on distance and throughput demands, transceivers might combine multiple channels to form a densely packed superchannel [1, 14], or divide each channel into multiple sub-carriers to ease DSP

and gain tolerance to fiber nonlinearities [15]. This broad range of operating conditions requires a powerful, highly flexible DSP [13]. A complete coherent DSP consists of multiple algorithms to compensate for both static and dynamic impairments [2]. These DSP algorithms can roughly be divided into two key categories; blind and pilot-aided algorithms [16]. Blind algorithms use knowledge of signal statistics such as modulation format and pulse shape. Using statistical estimates, the DSP algorithms are tuned to maximize performance. In contrast, pilot-based algorithms depend on knowledge of a fraction of the transmitted symbols (the pilots). This avoids the need for statistical estimates as the transmitted information is known. Pilot-aided algorithms are therefore inherently more resilient to noise, but the transmission of pilots implies a direct loss in throughput.

Early coherent transceivers transmitted polarization multiplexed quadrature amplitude modulation (PM-QPSK). Working with a constant-amplitude modulation format, blind algorithms such as the constant modulus algorithm (CMA) [17] and Viterbi-Viterbi (V-V) [18] phase tracking were used. Pilot-based DSP was also considered in this case to improve robustness [19]. As transceiver electronics evolved, higher-order formats such as PM M-ary quadrature amplitude modulation (PM-MQAM) were introduced to improve SE. These formats are inherently more sensitive to noise and require a higher signal-to-noise ratio (SNR) to achieve the same symbol error rate (SER) [1]. Moreover, since these formats carry more bits per information symbol, the SER can be considerably higher than the resulting BER. Blind DSP algorithms for higher-order PM-MQAM relies heavily on decision-directed (DD) processing and performance is therefore usually degraded when the SER increases. This issue has inherently become more cumbersome with the introduction of soft-decision (SD) FEC working at pre-FEC BERs exceeding 10^{-2} at which, assuming gray-coded signals, the corresponding SER can easily be > 10%. These errors in DD processing can be avoided using pilots but the key questions then are what is the needed pilot ratio and how it shall be optimized for different channel conditions. Previous work used OHs ranging from about 4% to 10% in the case of fully pilot-based DSP and higher order formats [20–22]. For only pilot-based CPE, OHs of about 1-2% have been reported, depending on implementation, laser linewidth and transmission distance [23, 24].

In this work we present a novel system-level pilot overhead optimization based on our pilot based DSP [25] and show detailed numerical and experimental studies of the implication under different operating conditions. We present a detailed comparison of the individual pilot-based algorithms with standard blind algorithms using Monte-Carlo simulations at target BER $4 \cdot 10^{-2}$. The simulations provide essential insights into how the required overhead for individual pilot-based algorithms contributes to the total overhead for fully pilot-based processing. This also provides new insights into the design trade-offs and the system-level differences between traditional blind and fully pilot-based DSP. To find the optimal total overhead, we propose to optimize the system achievable information rate (AIR) calculated from the generalized mutual information (GMI) in order to find the optimal trade-off between performance and the amount of inserted pilots. We transmit a 51×24 Gbaud PM-64QAM superchannel over 80 km spans of standard SMF (SSMF) with EDFA-only amplification using a recirculating loop. Comparing the optimum in back-to-back (B2B) with the optima after transmission over distances up to 1000 km, we find that pilot DSP is very robust to distortions from transmission and that the optimal overhead only weakly depends on transmission distance. The proposed pilot-DSP enables a spectral efficiency of 9.3 (8.3) bits/s/Hz after 480 (960) km of transmission. This corresponds to a superchannel throughput of 11.9 (10.6) Tb/s, demonstrating that low-overhead pilot-based DSP enables high performance flexible rate-reach transmission systems.

2. Pilot-based DSP

For the pilot-based DSP implementation we rely on transmission of frames. While normally not considered for blind DSP implementations we note that frames are already present in the

upper communication layers as well as needed for the FEC implementation so enforcing a framed structure on the DSP level does not drastically change the system design. However, joint design of DSP and FEC frames is beyond the scope of this work and we therefore only focus on the DSP frame design. The frame structure used here is illustrated in Fig. 1(a). Each frame consists of an initial pilot sequence followed by the data payload with periodically inserted pilots. The pilots are chosen from a single amplitude modulation alphabet such as M-ary phase shift keying (M-PSK). More specifically, in this work, all considered pilot symbols are PM-QPSK. The average energy is also chosen to be the same for both pilot and payload symbols, noting that this could be optimized to minimize the power penalty from pilot transmission. However, for implementation simplicity combined with low-overhead processing, we here use the same average energy for both pilot and payload symbols.
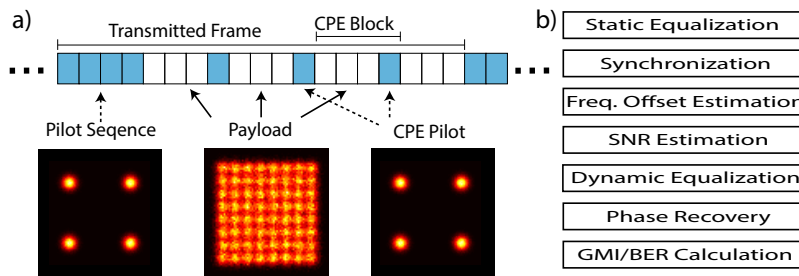

Fig. 1. (a) Schematic over the frame structured used to implement the pilot-based DSP. The frame consists of an initial sequence of pilots followed by the payload with periodically inserted, individual, pilot symbols for continuous phase tracking. The pilot symbols are PM-QPSK and the payload can be any arbitrary modulation format. The constellation diagrams show pilots and payload (PM-64QAM) for an SNR of 16.7 dB, corresponding to a payload BER of about $4 \cdot 10^{-2}$. (b) The order of the various DSP algorithms used to process the frame and extract the information carried by the payload symbols.

The DSP chain used to process the frame is shown in Fig. 1(b). It differs slightly from a traditional chain for blind DSP [2] and is instead designed together with the frame structure to ease the complexity and maximize the performance of each individual pilot algorithm. The first step in the DSP chain is static filtering to filter out out-of-band noise and compensate any dispersion. Following this, the synchronization stage is used to locate the starting position of the frame in the received signal. The order of the synchronization and static filtering steps can also be interchanged. Once located, the pilot sequence is used to find the frequency offset, estimate the signal SNR and set the dynamic equalizer to invert the channel response and fulfill the matched filtering criterion. Thus the adaptive filter taps are only updated at the beginning of the frame and are then kept constant until the next frame. The length of the frame therefore needs to be optimized with respect to the time scale of dynamic changes in the transmission channel (such as polarization rotations) and the OH from transmitting excess number of pilots. However, as discussed in section 3, the frame length in this work was limited by experimental constraints rather than dynamic channel fluctuations. Clock recovery can also be implemented using information from the pilot sequence. Following the pilot sequence, periodic pilot symbols are inserted into the payload and used for CPE only.

With the frame design using the sequence and periodic CPE pilots, we emphasize that the sequence length does not have to be constant for all frames. While challenging to meet in lab condition following limited available memory in the emulated transmitter and receiver, a practical implementation could first use an initialization phase with a long sequence and then switch transmission operation using a shorter sequence to simply correct the changes with respect to the last sequence. This can be directly understood using the dynamic equalizer as an example. In

this case, the adaptation for frame two is naturally performed using the found inverse channel response at the time occasion of the first sequence, in contrast to the first time the signal is synchronized when the adaptation has to start without any knowledge of the channel.

Figure 1(a) shows simulated constellation diagrams for a frame with PM-QPSK pilot symbols and PM-64QAM payload at a SNR of 16.7 dB (neglecting all impairments but AWGN). At this rate, we directly observe a key aspect of pilot-based processing, namely the effective SER difference between the pilots and the payload at the same SNR. In this example the SER of the 64QAM payload is 22%, while it is $< 10^{-4}$ for the pilots and thus significantly less averaging is needed compared to blind algorithms working on the payload symbols, as discussed below. In the following sections, we explain the details of the various algorithms in Fig. 1(b). For selected key cases, we also compare the performance of the proposed implementation with standard blind algorithms. Monte-Carlo simulations with 1000 realizations for each evaluated scenario were used to evaluated the performance considering QPSK, 16QAM, 64QAM and 256QAM payload symbols. In all cases, the simulations were performed with a SNR to obtain a BER of about $4 \cdot 10^{-2}$. The SNR values are then 5, 11.2, 16.7 and 22.1 dB for QPSK, 16QAM, 64QAM and 256QAM, respectively. This BER target correspond to about 20% FEC overhead, which has been reported when using state-of-the-art LDPC codes in [26]. For the pilot-case, we focus on the case with 16QAM payload to work with the lowest SNR for the pilots. For the case of frequency offset estimation (FOE) and carrier phase recovery (CPE), we have considered a target symbol rate of 24 Gbaud. When evaluating the CPE, each realization consisted of 2^{17} complex symbols.

2.1. Static equalization

The static equalizer is used to compensate chromatic dispersion and in this work it is assumed that the link length is known. However, multiple methods to estimate the amount of dispersion are available [27, 28]. In this work we use the static equalizer prior to the frame synchronization. While we note that this is not necessary, it helps to improve the robustness. The static equalizer can also contain a fixed receiver-side matched filter [2] (such as the second root-raised cosine filter (RRC)) and filters to compensate known bandwidth limitations from pre-characterized components such as DACs, ADCs and the modulator [29]. Depending on the number of taps needed for each filter, they can be implemented in either time or frequency domain and multiple filters can be combined into a single filtering stage. For dispersion, time domain filtering is usually more effective for distances below 150 km [30] whereas frequency-domain implementation reduces computational complexity for long-haul applications [31].

However, as the static filtering is performed prior to correcting the frequency offset, careful design is needed to ensure maximal performance. For a static dispersion compensating equalizer, a frequency offset will result in a residual delay. However,. the syncing stage will align the frame and this delay should not affect the DSP performance. A static filter used to compensate non-perfect receiver components should be placed prior to the frequency offset estimation (or being frequency shifted accordingly), as this filter should compensate the signal as it was measured. In direct contrast, a receiver-side matched filter should be placed after as it otherwise will filter out part of the signal and therefore degrade the SNR, depending on the exact offset [2]. If all filters are used, the static equalization filter is preferably separated into two filters with one placed before and one after the frequency offset estimation block.

2.2. Synchronization

The synchronization is required to identify the starting position of each frame, a core part which is not needed for blind processing. In our DSP, synchronization is performed on a sample level. Sub-sample timing is handled by the adaptive equalizer (see Section 2.5) or using a separate timing recovery algorithm [2]. The synchronization implemented here consists of 2 parts to ease computation, an initial coarse stage (which is optional but used to improve speed and robustness)

followed by a fine alignment.

The coarse synchronization uses CMA with adaptive step size to improve convergence speed. The signal is divided into N segments and each frame is processed individually in parallel using CMA. Exploiting the fact that the CMA error function is ideal for M-PSK, the segment achieving the smallest residual error is selected. While this stage naturally performs the best if the relative difference in modulation order is large, we have found it to work well in experiments even with PM-8QAM payload. Important to note here is that we used significantly less taps for the synchronization compared to the actual dynamic equalizer (see Section 2.5). In experiments, we found that about 15 taps is typically enough for synchronization. In addition we note that the proposed coarse synchronization technique relies on the payload being modulated using a multi-modulus modulation format with the pilots being a constant modulus format. For QPSK pilot symbols we found that having PM-8QAM payload did not pose any problem for the proposed synchronization technique neither in experiments nor simulations. If the payload also consists of a constant modulus modulation format, this stage can be omitted and the synchronization can be entirely performed using an initial CMA-based equalization followed by fine alignment using cross-correlation. However, with the focus of this work being on DSP for high-order modulation formats we use the proposed coarse synchronization throughout this work.

Using the symbols found and processed in the coarse stage, we perform a coarse frequency offset estimation to compensate the linear part of the phase evolution. A cross-correlation-based estimation is then used for final adjustment. The cross-correlation uses the phase of the sent pilot symbols and the selected output symbols from the coarse synchronization to find the final difference. Given the large tolerance to phase noise from QPSK, we did not include any CPE (see Section 2.6) in the synchronization stage.

While doing the coarse estimation only for polarization is sufficient, we do the cross-correlation-based alignment for every mode present in the system to allow for accurate alignment even in the presence of strong polarization mode dispersion (PMD) or differential mode-group delay (for systems transmitting information on multiple spatial modes). Finally, we also note that while often neglected in experiments using blind DSP, some kind of reference is required when using PM-MQAM formats. First of all, the receiver side cannot distinguish between the X-polarization and the Y-polarization states since they have the same statistics. Moreover, the complex modulation formats used are typically $\pi/2$-symmetric and therefore indistinguishable to such rotations [32]. Due to this ambiguity, without any reference of the absolute phase, the phase cannot be aligned and successful transmission can therefore not be assumed when using standard PM-MQAM formats and normal bit encoding, regardless of how accurate the CPE algorithm itself is. One way of overcoming this is to use differential encoding but this comes with the price of reduced information rate and requires special FEC design [33].

2.3. Frequency offset estimation

A frequency offset arises from a non-zero difference in frequencies between the transmitter laser and the local oscillator. The range of the offset will depend on the kind of laser used but considering standard external cavity lasers, the frequency offset can reach a few GHz [34]. On a symbol level, this gives rise to a linear phase increase between consecutive symbols according to

$$y_k = x_k \exp\left(j2\pi k\Delta f T_0\right) + n_k, \tag{1}$$

with x_k denoting the transmitted symbol at time instance k after matched filtering and down sampling, y_k the received symbol, Δf the frequency offset, T_0 the symbol period and n_k circular symmetric AWGN. Assuming the use of complex modulation formats, FOE is challenging as data is encoded on the phase, causing it to fluctuate with the random data. This increases the required averaging length, making it more challenging in regions with lower SNR.
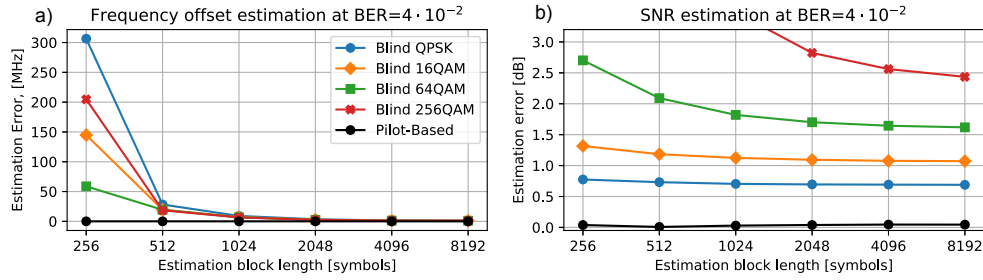
Fig. 2. Blind vs. pilot-aided (a) frequency offset and (b) SNR estimation for different estimation block lengths. The algorithmic implementations are described in Section 2.3 and Section 2.4, respectively.

For systems without pilots, blind estimation algorithms usually rely on trying to remove the phase modulation using Viterbi-Viterbi schemes such as raising a QAM constellation to the 4th power and seek for the linear phase offset. While this works well for single amplitude formats such as QPSK, the data modulation cannot be fully removed for higher order QAMs using a single operation and the estimation quality is therefore degraded. A popular way of implementing the search for a linear frequency offset is to use a Fourier transform [35]. Similarly, FOE can be realized by Fourier transforming the spectrum and compare offsets on both sides surrounding DC. The issue with both these methods is the resolution as practical fast Fourier transform (FFT) sizes are dictated by hardware constraints and large FFT size causes excessive power dissipation and the goal is therefore to use the minimal possible size. Considering real-time frequency domain implementation of dispersion compensation, a 512 sample (2-folded oversampling) FFT can compensate up to 3500 km for 12.5 Gbaud [31]. In general, it is preferred to have all FFT block sizes being smaller than the dispersion compensation block so that all operations can be implemented using a single FFT.

In contrast, FOE can be done very effectively using the pilot sequence. Removing data modulation is trivial as the symbols are known to the receiver, avoiding the issues of 4:th power implementations. The phase for symbol k can be written as

$$\angle y_k = 2\pi k f_0 \angle x_k + \epsilon_k, \tag{2}$$

with f_0 denoting the frequency offset and ϵ_k any distortion (including both AWGN and phase noise). Replacing k with vector notation, the least mean squares (LMS) fit to f_0 can be found as by simply calculating the expectation value of the unwrapped phase difference between consecutive symbols after data modulation.

The results of frequency offset estimation are shown in Fig. 2(a), comparing the blind 4:th-power method using implemented using FFT with pilot-based estimation. Assuming f_0 being uniformly distributed between 0 and 2 GHz, we observe that pilot-based FOE, even with 16QAM payload, achieves negligible frequency error even when only 256 symbols are used, in contrast to the blind method which requires 2048 symbols to achieve similar performance. In addition, we also note that the range for the blind method is limited to $\pm R_S/8$ assuming two-folded oversampling with R_S denoting the symbol rate. The pilot-based FOE is, however, only limited by its oversampling ratio.

2.4. SNR estimation

Estimating the SNR is crucial to ensure maximum performance of optical transceivers. Considering modern systems using SD-FEC, the input to the decoder consists of log-likelihood

ratios (LLRs) which has to be calculated to enable decoding. Assuming an AWGN channel with a bit-wise memory-less receiver and uniform symbol probabilities, the LLRs are calculated according to [36]

$$L_i = \log \frac{\sum_{x \in \chi_i^1} \exp\left(-\rho|y-x|^2\right)}{\sum_{x \in \chi_i^0} \exp\left(-\rho|y-x|^2\right)}, \tag{3}$$

where $i \in 1, 2, \ldots \log_2(M)$ the considered bit, $\chi_i^{1/0}$ the symbols mapping bit i to 1/0, respectively, and ρ is the SNR. Note that for transmitters using probabilistic shaping (PS), Eq. 3 needs to be modified to account for a non-uniform symbol distribution. In addition, the optimal shaping parameter is SNR dependent and accurate estimation of the SNR is therefore crucial to maximize the performance of PS.

There are multiple ways of estimating the SNR, for Gaussian channels. While the fiber channel is not a strict Gaussian channel, we here focus on uncompensated transmission in which nonlinear signal distortions can be accurately modelled as an additional AWGN contribution [9]. Defining the SNR as E_s/N_0 with E_s denoting the signal power and N_0 the noise variance, we directly observe that in order to estimate the SNR from received symbols, the modulation (both amplitude and phase) has to be removed. SNR estimators are therefore typically divided into pilot-aided estimators with knowledge of the transmitted symbol x and blind estimators which relies on estimating x from knowledge of the signal statistics.

A comparison between blind estimation based on DD processing and pilot-aided SNR estimation is shown in Fig. 2(b). We observe a large penalty for blind estimation, especially for large M. This follows the trend of estimation accuracy which directly depends on the SER and the performance therefore degrades for higher order formats which can have significantly higher SER than BER. In contrast, we note that the pilot-based estimation produces an estimation error which is <0.1 dB. In addition, when normalizing the signal to calculate the reference points, the unknown noise level causes an offset which depends on the actual SNR value. To overcome this issue, iterative re-normalization using the estimated SNR to subtract the noise power or using estimated reference points rather than the actual transmitted one should be used. However, this makes the estimation more complex.

2.5. Dynamic equalization

Following the static equalizer, a dynamic equalizer is used to implement polarization de-multiplexing, compensate for residual chromatic dispersion and mitigate any unknown filtering penalties to reach the matched-filtering criterion before down-sampling the signal to 1 sample per symbol. Multiple pilot-aided estimation methods exist to find the inverse channel response [37]. Here, we have used a conventional time-domain filter with dynamic coefficients. The coefficients were updated with a standard LMS algorithm using the pilot-knowledge to calculate the error. In addition, we use an adaptive step size to ensure rapid convergence [38].

In contrast to the CMA-based processing of the coarse synchronization stage, the dynamic equalizer uses more taps to cover the complete temporal duration of the signal. For processing the experimental data (see Section 4), we found that 45 T/2-spaced taps were sufficient for 24 Gbaud signals with 1% roll-off. While the frequency offset has been removed prior to this stage (see Fig. 1(b), we did not include any CPE in the LMS-based estimation following the short pilot sequence length and the very accurate pilot-based FOE (see Section 2.3). While this theoretically is needed for the LMS algorithm, we found that the impact of phase noise on the PM-QPSK pilot sequence was very minor and did not affect the pilot-aided equalizer convergence. This was verified using both the experimental data and numerical simulation emulating a maximum per-laser linewidth of 100 kHz.

The update requirements of the dynamic equalizer depends, to a large degree, on the speed of the polarization fluctuations occurring in the fiber. Following extreme changes in the surrounding

environment, such as lightening strikes, the state of polarization (SOP) can fluctuate very fast [39]. However, long-term measurements of SOP-changes in installed fiber indicates much slower changes [40]. With our approach using a pilot sequence to estimate the channel and update the dynamic taps, it is clear that the maximum total frame length will depend strongly on the surrounding environment as a too long frame will induce additional penalties from SOP tracking errors. A pure polarization rotation, corresponding to the worse-case rotation along S2 in Stokes representation, can be modelled using the Jones representation according to

$$
\begin{pmatrix} \hat{x}_{k_x} \\ \hat{x}_{k_y} \end{pmatrix} = \begin{pmatrix} \cos(\Theta) & -\sin(\Theta) \\ \sin(\Theta) & \cos(\Theta)) \end{pmatrix} \begin{pmatrix} x_{k_x} \\ x_{k_y,} \end{pmatrix}
\tag{4}
$$

where subsubscript x/y denotes the x and y polarization, respectively, ˆ denotes the rotated output SOP and Θ the rotation angle. Figure 3(a) shows the GMI reduction caused by residual angle miss match for QPSK, 16QAM, 64QAM and 256QAM, respectively. As expected, we observe a penalty from a non-ideal polarization demultiplexing stage which grows rapidly with modulation order. Using additional numerical simulation we found, however, that the pilot-based equalizer was capable of detecting the rotation angle with a negligible error at the considered SNRs, independent of the angle Θ and any applied PMD, as long as the differential group delay was within the equalizer memory. In order to investigate how the optimal frame length depends on a linear polarization rotation occurring between consecutive pilot sequences we simulated a continuous linear polarization rotation. We consider a rotation speed corresponding to a total rotation Θ occurring over 2^{18} symbols. The pilot sequence length is set to 2048 symbols (as was found to be optimal in the experimental evaluation, see Section 4) and the resulting AIR for frame lengths between 2^{14} and 2^{18} symbols is shown in Fig. 3(b). As further discussed in Section 3 we found that the the SOP was stable enough so that the frame length was limited by experimental constraints rather than environmental fluctuations. Considering the largest investigated rotation $\Theta = 10^o$, this corresponds to a rotation speed of about 16 krad/s for a 24 Gbaud signal. While this tracking speed seems to be sufficient even for SSMF in shaky environments [41], we note that detailed studies of this will require active field trials, far beyond the scope of this work. If large tracking speed is required, non-uniform pilot sequence lengths or updates using the phase pilots can be used to increase speed while maintaining low equalization OH as discussed in Section 2.
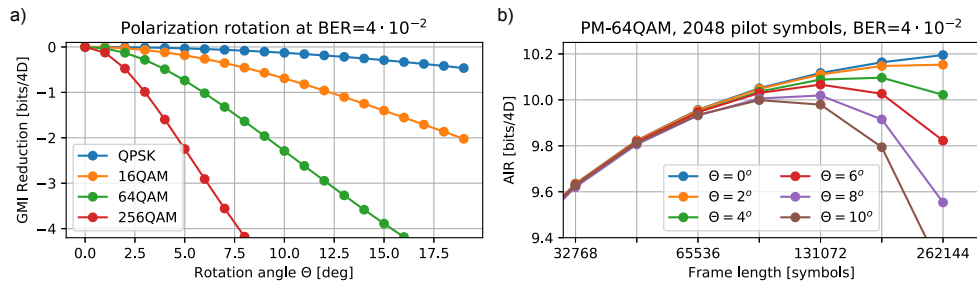

Fig. 3. Simulation of sensitivity to polarization rotations caused by dynamic environmental fluctuations in the fiber. (a) GMI reduction (penalty) for a rotation of Θ degrees. (b) AIR optimization considering a pilot sequence of 2048 symbols. The rotation Θ here occurs over a total time corresponding to 2^{18} symbols. As such, a rotation angle $\Theta = 10^o$ corresponds to a rotation speed of about about 16krad/s considering a 24 Gbaud signal.

2.6. Carrier phase estimation

Pilot-based CPE offers a powerful alternative to blind processing, which also mitigates the risk for cycle slips [42, 43]. Pilot-based CPE relies on interpolating between the periodically inserted pilots. CPE can be done using either the pilots only or in a hybrid configuration in which the pilots do a coarse tracking and ensure no cycle slips and a blind CPE algorithm is used for tracking any residual offset [44]. However, here we focus on a fully pilot-based CPE. As the phase noise can be modelled as a Brownian motion [45], the precise random walk can not be tracked using the simplest form of piece-wise continuous linear interpolation. While more sophisticated interpolations can be done using Kalman filters [46], we here restrict ourselves to the simplest case as Kalman filters are recursive and therefore cannot be parallelized for hardware implementation of DSP.
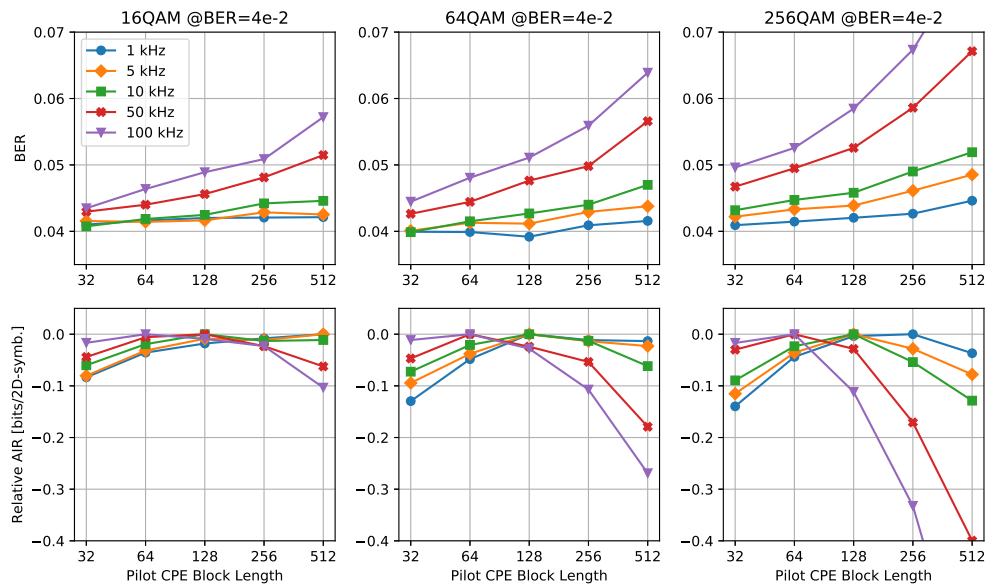

Fig. 4. Performance of pilot-based CPE using simple linear interpolation for 16QAM, 64QAM and 256QAM at a target BER of $4 \cdot 10^{-2}$. The overall simulated linewidth is a factor of two larger, accounting for equivalent laser performance in both the transmitter and the receiver. The AIR is calculated by deducting the CPE pilot OH from the measured GMI and each curve is normalized to its maximum value.

The performance dependence on pilot OH is shown in Fig. 4 using both BER and AIR. We investigate laser linewidths of 1, 5, 10, 50 and 100 kHz and assume that the combined linewidth of both the transmitter and receiver laser is a factor of two larger. The AIR values are calculated by deducting the pilot OH from the measured GMI and normalizing to the maximum value for each considered linewidth. The increased sensitivity with respect of format cardinality is clearly visible in Fig. 4. For 16QAM, we do not observe a strong dependence on the found optima for different linewidth but the expected parabolas can clearly be seen for linewidths exceeding 10 kHz. Following the trend of increased sensitivity, a large and rapidly growing penalty is observed for 256QAM if the linewidth exceeds 10 kHz. For PM-64QAM we observe a results which, naturally, lies in between the case of PM-16QAM and PM-256QAM. For linewidths up to 50 kHz, the AIR penalty is below 0.1 bits/2D-symbol for block-lengths between 64 and 512 symbols. Looking at the BER results, we observe the same trend for all modulation formats. Important to note is also the increased sensitivity to linewidth compared to what was previously reported in [47], owing to the additional noise present when considering SD-FEC [48].

3. Experimental setup

While the performance of individual pilot-based algorithms was evaluated in Section 2, the actual performance of an algorithm depends on the input signal quality together with the performance of the algorithms used prior in the DSP-chain. For example, a non-converged equalizer will drastically reduce the SNR and any residual frequency offset will have to be tracked by the CPE. The later effect can be modelled as a biased random walk and the required performance of the CPE can differ a lot when accounting for a small, MHz-level, residual frequency offset [48]. In order to evaluate the DSP performance, full system measurements are therefore needed. The experimental setup used to do this is shown in Fig. 5. We used a standard external cavity laser (ECL) (100 kHz specified maximum linewidth) centered on 1545.3 nm to seed an electro-optic (EO) frequency comb generating about 50 lines with 25 GHz spacing. The EO-comb was built using two phase modulators cascaded with an intensity modulator for flatness, similar to [49]. The comb output was amplified, before being flattened and filtered using a wavelength selective switch (WSS) to select 51 lines.
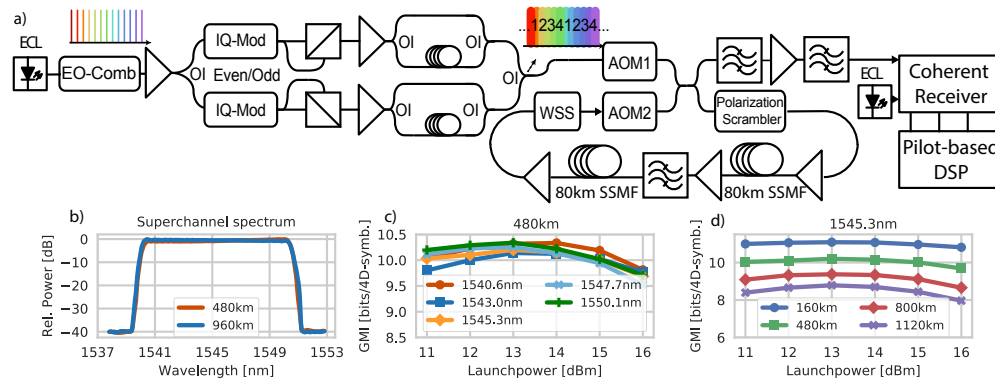

Fig. 5. (a) Experimental setup used for pilot-DSP evaluation in transmission of the 51×24 Gbaud PM-64QAM superchannel. (b) Measured superchannel spectrum after 480 km and 960 km of transmission, respectively. (c) Launchpower sweep for 5 selected evenly spaced test channels after 480 km. (d) Launch power sweep at varying distance for center channel at 1545.3 nm.

Following amplification, a 25 GHz optical interleaver (OI) was used to divide the lines into even and odd lines. These were modulated independently using two IQ-modulators driven by two 60 GS/s DACs to modulate the 24 Gbaud PM-64QAM signal shaped with a root-raised cosine filter with 1% roll-off. The maximum frame-length available was limited to 104704 symbols by the memory of the memory board controlling the DACs. After modulation, a PM signals were emulated using the split-delay-combine method with delay of about 250 symbols. To reduce the non-linear penalty from having correlated data [50], two sets of 50 GHz OIs were used to delay every second even/odd channel with an additional delay of about 750 symbols, creating a 123412... decorrelation scheme. The even and odd channels were then combined and sent to the recirculating loop setup.

The recirculating loop consisted of two spans of 80 km SSMF with about 16 dB loss each. A fixed 11 nm band-pass filter, an additional WSS and a loop-synchronous polarization scrambler were also placed inside the loop. The WSS was used both to filter out-of-band amplifier noise and to flatten the gain tilt caused by the EDFAs (about 5.5 dB noise figure each). The measured spectrum after 480 km and 960 km can be seen in Fig. 5(b). At the receiver, the channel under test was selected using a cascade of an initial 0.3 nm filter, an EDFA and a final 0.8 nm filter. A free-running ECL, equivalent to the transmitter seed laser, and a 23 GHz analog bandwidth

optical hybrid was used to detect the optical signal. The resulting electrical signals were digitized using a 50 GS/s real-time oscilloscope with 23 GHz analog bandwidth. The samples were then processed offline using the DSP outlined in section 2. The DSP is also available online under a general public licence (GPL) [51].

4. Results

To evaluate the pilot-based DSP, we focused our study on PM-64QAM payload for distances up to about 1000 km. First, we performed launch power optimization to balance amplifier noise and non-linear penalty. The result for 5 evenly spaced test channels after 480 km is shown in Fig. 5(c). We observe a slight wavelength dependence on the optimal launch power for each test channel with a combined launch power of 13 dBm (-4 dBm/channel) resulting in the best performance. Launch power sweeps for multiple distances using the center channel is shown in Fig. 5(d). As expected, we observe an increased sensitivity to non-optimized launch powers at longer distances.

4.1. Pilot-overhead optimization

As previously outlined, optimization of the pilot OH is crucial in order to maximize the performance of systems using pilot-based DSP. We therefore varied both the length of the pilot sequence and the CPE pilot insertion ratio and measured the resulting GMI. We first investigated any penalty from polarization fluctuations and observed a constant measured GMI along the detected frames even at the longest possible (see Section 3) frame length of 104704 symbols. We therefore concluded that SOP fluctuations were small enough to not induce any further penalty and always used the longest possible pilot frame. The sequence OH was therefore only changed by changing the length of the sequence itself.
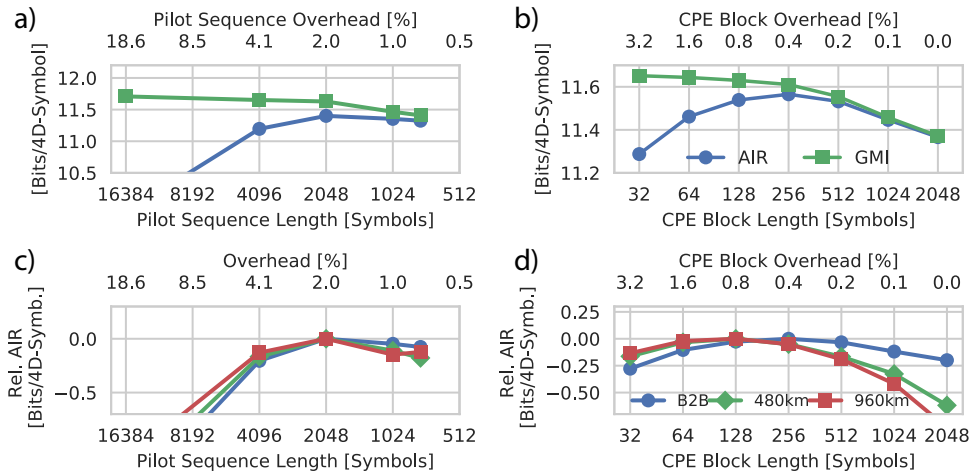

Fig. 6. Pilot overhead optimization using AIR. (a) Measured GMI and corresponding AIR for varying the pilot sequence length from 16384 to 768 symbols. (b) Corresponding results for varying the CPE pilot block length from 32 to 2048 symbols. (c) and (d) Distance dependence for the found optima for pilot sequence and CPE block length, respectively.

The results are shown in Fig. 6. The optimization was done by calculating the AIR from the measured GMI by deducting the pilot OH. A B2B comparison of both GMI and AIR as a function of the sequence length and the CPE block length are shown in Fig. 6(a) and 6(b), respectively. We observe that while the GMI increases with the OH, the AIR produces an optimum for both cases.

For the pilot sequence, the found optimum was 2048 symbols, corresponding to an overhead of 2%. Moreover, we observe a small performance difference with respect to 1024 symbols, highlighting the trade-off between increased performance and additional OH. We also note that the results in Fig. 6(a) shows the optimal sequence length for all algorithms working on the pilot sequence, confirming that no significant scaling in OH is observed when comparing individual algorithms to a complete processing chain. For the CPE block length, we find an optimal insertion ratio of 1/256, corresponding to 0.4% OH. In all cases, we optimized the CPE pilot averaging which, at optimal insertion ratio, was 4 symbols. In addition, comparing the measured GMI to the simulations shown in Fig. 4, we observe that the measured performance matches well with a simulated laser linewidth of 10 kHz, one order of magnitude below the specified maximum linewidth. This value matched well with extraction of a Lorentzian-like frequency noise floor from the frequency noise measurements of the ECL using coherent detection and spectral processing. Direct fitting of a Lorentzian envelope to the beating spectrum observed on a spectrum analyzer revealed higher linewidths, a difference which we attributes to the presence of 1/f and Gaussian components in the laser frequency noise [52]. We further note that while doing joint CPE over both polarization, the overhead can be reduced with about a factor of two [43]. However, following the polarization emulation stage used (see Section 3), independent control of the pilot distribution was not possible and we therefore used independent CPE for both polarizations.

Figures 6(c) and 6(d) compare the AIR for the center channel using various OH configurations at a transmission distance of 480 and 960 km, respectively. For the sequence, we observe that the found B2B optima remains optimal for all considered distances, highlighting the robustness of the pilot-based approach. A slight distance dependence can be observed for the CPE block lengths which is expected when accounting for additional noise from the fiber channel which requires slightly higher averaging to maintain optimal performance. However, we note that the difference between a CPE block length of 128 and 256 symbols was 0.05 bit/4D-symbol for after both 480 km and 960 km. Therefore, B2B optimization is sufficient to find target OHs for a broad range of distances. At longer distances, a slightly shorter CPE block length combined with a longer block averaging filter is preferable in order to add resilience to both additive noise, which requires a larger averaging window and therefore reduce the tracking speed, and additional non-linear phase noise from the fiber channel.
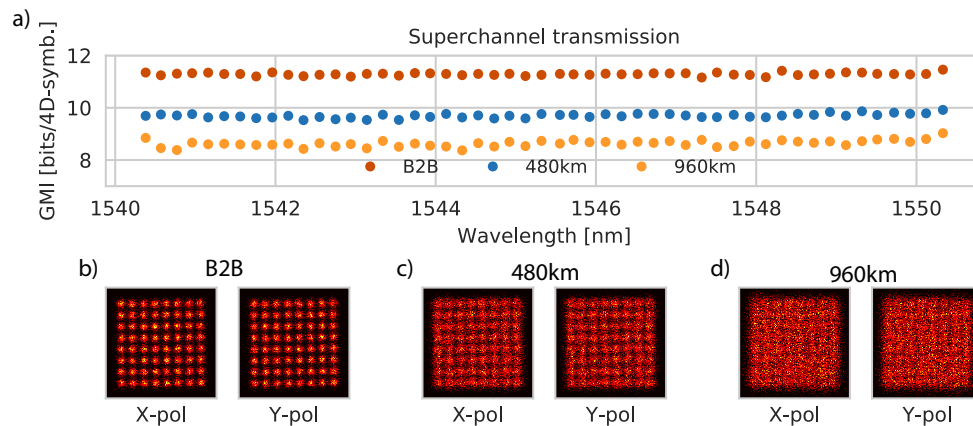
4.2. Long-haul transmission

Fig. 7. (a) Measured GMI for all 51 channels in B2B and after 480 and 960 km of transmission. (b) - (d), corresponding constellation diagrams for the center channel at 1545.3 nm.

Finally, to verify the capability of our proposed pilot-based DSP in high SE transmission, we performed a full superchannnel transmission using the optimized OHs in B2B from Section 4.1. The measured GMI (after deducting pilot OH) in B2B, after 480 km and after 960 km of transmission for all 51 channels is shown in Fig. 7(a). The GMI in each point was estimated from $>10^6$ bits per batch and averaged over 5 measurement batches. The GMI average over the superchannel was 11.3 bits/4D-symbol in B2B and 9.7 and 8.6 bits/4D-symbol after 480 and 960 km, respectively. The corresponding superchannel throughput in the two later cases, assuming optimized coding for each channel, was 11.9 Tb/s and 10.6 Tb/s, respectively.

Considering the 1 GHz inter-channel guard-band within the superchannel, corresponding to 4% OH, and the 2.4% pilot OH, the resulting SE after 480 and 960 km was 9.3 and 8.3 bits/s/Hz, respectively. Constellation diagrams for the center channel at the three distances, respectively, are shown in Figs. 7(b)-7(d). Finally we emphasize that no adaptation of any DSP algorithm was used when processing the 51 channels forming the superchannel. Finally, we note that the GMI-based performance estimation assumes completely independent received symbols. This implies that presence of any error bursts or memory effects have to be removed prior to decoding, using techniques such as interleaving [36], in order for our performance estimation to be valid.

5. Discussion

Pilot-based processing offer the advantage of a flexible and robust DSP which is inherently modulation format independent. Pilot-based DSP is therefore well suited for flexible format transceivers which should be able to rapidly adapt without requiring a full reconfiguration of the DSP [13]. From a system-level perspective, the pilot OH has to be compared to other OHs in the system. Using OH optimization, as shown in Section 4, we note that the optimal pilot OH of 2.4% is the smallest OH when calculating the SE from the measured GMI. Inter-channel guard-bands result in 4% loss in SE and considering coding OH of 20% and beyond, the pilot OH is very small.

To reduce the OH even further, techniques such as joint processing over the two polarizations [53] and hybrid approaches can be used. Joint processing is powerful and for modern SMF, the low PMD allows for reducing the OH with about a factor of two [43]. Hybrid pilot-blind approaches reduce the OH even further using techniques such as encoding information on the QPSK CPE pilots and use a standard V-V algorithm to track the phase. However, the gain of such schemes always has to be traded against the added complexity for a relative small gain in channel rate as the CPE pilot OH is very minor compared to e.g. coding OH.

Scaling modulation order while maintaining reasonable DSP complexity is very challenging when designing high SE transceivers. This can be directly understood by comparing various CPE schemes. A comparison of hardware implementations of BPS and pilot-based CPE demonstrated the clear benefit of using pilots for CPE [54]. As BPS relies on using a number of test angles to find the one giving the minimum resulting error, the number of angles required to maintain a given SER/BER increases with modulation order. Following this, the power consumption of BPS increases with modulation order, in contrast to pilot-based CPE which is virtually unaffected [54].

6. Conclusion

We have described and investigated fully pilot-based DSP focusing on advanced modulation formats and powerful forward error correction. Using numerical simulation, we have evaluated the performance of individual algorithms building up the DSP chain and compared to powerful blind algorithms. The pilot-based algorithms show a large resilience to noise and distortions and operations such as correcting frequency errors and estimating the SNR can be performed with significantly less symbols compared to standard blind approaches. In addition, we demonstrate transmission of a 51×24 Gbaud PM-64QAM superchannel over distances reaching 1000 km. We use achievable information rate to optimize the pilot overhead in order to maximize the

resulting spectral efficiency. We find the that optimum overhead only weakly depends on transmission distance, enabling the optimum found in back-to-back configuration to be used for all considered distances. Following this, the optimized DSP enabled a spectral efficiency of 9.3 (8.3) bits/s/Hz, corresponding to 11.9 (10.6) Tb/s superchannel throughput, after 480 (960) km of transmission. Our results show the feasibility for pilot-based DSP to enable transmission of high spectral efficiency channels with flexible target distance and modulation format by allowing stable operation despite large changes in received signal to noise ratio.

Funding

References

1. P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: the previous 20 and the next 20 years [invited]," Opt. Express **26**, 24190–24239 (2018).
2. M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," J. Light. Technol. **35**, 1125–1141 (2017).
3. A. Leven and L. Schmalen, "Status and recent advances on forward error correction technologies for lightwave systems," J. Light. Technol. **32**, 2735–2750 (2014).
4. C. Laperle and M. Osullivan, "Advances in high-speed DACs, ADCs, and DSP for optical coherent transceivers," J. Light. Technol. **32**, 629–643 (2014).
5. S. Varughese, J. Langston, V. A. Thomas, S. Tibuleac, and S. E. Ralph, "Frequency dependent ENoB requirements for M-QAM optical links: An analysis using an improved digital to analog converter model," J. Light. Technol. **36**, 4082–4089 (2018).
6. A. Napoli, M. Mezghanni, T. Rahman, D. Rafique, R. Palmer, B. Spinnler, S. Calabro, C. Carlos, M. Kuschnerov, and M. Bohn, "Digital compensation of bandwidth limitations for high-speed DACs and ADCs," J. Light. Technol. **6**, 1 (2016).
7. T. Duthel, P. Hermann, J. Scheiss, C. R. S. Fludger, A. Bishplinghoff, and T. Kupfer, "Characterization and pre-distortion of linear and non-linear transmitter impairments for PM-64QAM applications," in *European Conference on Optical Communication (ECOC),* (2016), p. W.4.P1.SC3.25.
8. R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," Opt. Express **21**, 25685–25699 (2013).
9. P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," J. Light. Technol. **32**, 694–721 (2014).
10. E. Agrell, M. Karlsson, A. R. Chraplyvy, D. J. Richardson, P. M. Krummrich, P. Winzer, K. Roberts, J. K. Fischer, S. J. Savory, B. J. Eggleton, M. Secondini, F. R. Kschischang, A. Lord, J. Prat, I. Tomkos, J. E. Bowers, S. Srinivasan, M. Brandt-Pearce, and N. Gisin, "Roadmap of optical communications," J. Opt. **18**, 063002 (2016).
11. A. Ghazisaeidi, L. Schmalen, I. F. de Jauregui Ruiz, P. Tran, C. Simonneau, P. Brindel, and G. Charlet, "Transoceanic transmission systems using adaptive multirate FECs," J. Light. Technol. **33**, 1479–1487 (2015).
12. T. Fehenberger, A. Alvarado, G. Böcherer, and N. Hanik, "On probabilistic shaping of quadrature amplitude modulation for the nonlinear fiber channel," J. Light. Technol. **34**, 5063–5073 (2016).
13. G. Bosco, "Advanced modulation techniques for flexible optical transceivers: The rate/reach tradeoff," J. Light. Technol. **37**, 36–49 (2019).
14. X. Liu, S. Chandrasekhar, and P. J. Winzer, "Digital signal processing techniques enabling multi-Tb/s superchannel transmission: An overview of recent advances in DSP-enabled superchannels," IEEE Signal Process. Mag. **31**, 16–24 (2014).
15. F. P. Guiomar, A. Carena, G. Bosco, L. Bertignono, A. Nespola, and P. Poggiolini, "Nonlinear mitigation on subcarrier-multiplexed PM-16QAM optical systems," Opt. Express **25**, 4298–4311 (2017).
16. M. Kuschnerov, M. Chouayakh, K. Piyawanno, B. Spinnler, E. de Man, P. Kainzmaier, M. S. Alfiad, A. Napoli, and B. Lankl, "Data-aided versus blind single-carrier coherent receivers," IEEE Photonics J. **2**, 387–403 (2010).
17. A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," IEEE Transactions on Signal Process. **44**, 1136–1155 (1996).
18. A. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," IEEE Transactions on Inf. Theory **29**, 543–551 (1983).
19. R. Elschner, F. Frey, C. Meuer, J. K. Fischer, S. Alreesh, C. Schmidt-Langhorst, L. Molle, T. Tanimura, and C. Schubert, "Experimental demonstration of a format-flexible single-carrier coherent receiver using data-aided digital signal processing," Opt. Express **20**, 6 (2012).
20. T. Rahman, D. Rafique, B. Spinnler, S. Calabr, E. D. Man, A. Napoli, M. Bohn, G. Khanna, N. Hanik, S. Member, E. Pincemin, C. L. Bou, S. Bordais, C. Andr, C. Dourthe, and B. Ragu, "Long-haul transmission of PM-16QAM, PM-32QAM and PM-64QAM based terabit superchannels over a field deployed legacy fiber," J. Light. Technol. **34**, 3071–3079 (2016).

21. M. P. Yankov, E. P. da Silva, F. D. Ros, and D. Zibar, "Experimental analysis of pilot-based equalization for probabilistically shaped WDM systems with 256QAM/1024QAM," in *Optical Fiber Communication Conference,* (Optical Society of America, 2017), p. W2A.48.

22. H. Hu, M. P. Yankov, F. D. Ros, Y. Amma, Y. Sasaki, T. Mizuno, Y. Miyamoto, M. Galili, S. Forchhammer, L. K. Oxenlowe, and T. Morioka, "Ultrahigh-spectral-efficiency WDM/SDM transmission using PDM-1024-QAM probabilistic shaping with adaptive rate," J. Light. Technol. **36**, 1304–1308 (2018).

23. M. Pajovic, D. S. Millar, T. Koike-Akino, R. Maher, D. Lavery, A. Alvarado, M. Paskov, K. Kojima, K. Parsons, B. C. Thomsen, S. J. Savory, and P. Bayvel, "Experimental demonstration of multi-pilot aided carrier phase estimation for DP-64QAM and DP-256QAM," in *European Conference on Optical Communication (ECOC),* (2015), p. Mo.3.3.2.

24. J.-X. Cai, H. G. Batshon, M. V. Mazurczyk, O. V. Sinkin, D. Wang, M. Paskov, W. W. Patterson, C. R. Davidson, P. C. Corbett, G. M. Wolter, T. E. Hammon, M. A. Bolshtyansky, D. G. Foursa, and A. N. Pilipetskii, "70.46 Tb/s over 7,600 km and 71.65 Tb/s over 6,970 km transmission in C+L band using coded modulation with hybrid constellation shaping and nonlinearity compensation," J. Light. Technol. **36**, 114–121 (2018).

25. M. Mazur, J. Schröder, A. Lorences-Riesgo, M. Karlsson, and P. A. Andrekson, "Optimization of low-complexity pilot-based DSP for high spectral efficiency 51×24 Gbaud PM-64QAM transmission," in *European Conference on Optical Communication (ECOC),* (2018), p. MoF4.2.

26. F. Buchali, A. Klekamp, L. Schmalen, and T. Drenski, "Implementation of 64QAM at 42.66 GBaud using 1.5 samples per symbol DAC and demonstration of up to 300 km fiber transmission," in *Optical Fiber Communication Conference,* (Optical Society of America, 2014), p. M2A.1.

27. H. Wymeersch and P. Johannisson, "Maximum-likelihood-based blind dispersion estimation for coherent optical communication," J. Light. Technol. **30**, 2976–2982 (2012).

28. S. Yao, T. A. Eriksson, S. Fu, P. Johannisson, M. Karlsson, P. A. Andrekson, T. Ming, and D. Liu, "Fast and robust chromatic dispersion estimation based on temporal auto-correlation after digital spectrum superposition," Opt. Express **23**, 15418 (2015).

29. D. Rafique, A. Napoli, S. Calabro, and B. Spinnler, "Digital preemphasis in optical communication systems: On the DAC requirements for terabit transmission applications," J. Light. Technol. **32**, 3247–3256 (2014).

30. C. Fougstedt, A. Sheikh, P. Johannisson, A. G. i Amat, and P. Larsson-Edefors, "Power-efficient time-domain dispersion compensation using optimized fir filter implementation," in *Advanced Photonics 2015,* (Optical Society of America, 2015), p. SpT3D.3.

31. D. E. Crivelli, M. R. Hueda, H. S. Carrer, M. del Barco, R. R. Lopez, P. Gianni, J. Finochietto, N. Swenson, P. Voois, and O. E. Agazzi, "Architecture of a single-chip 50 Gb/s DP-QPSK/BPSK transceiver with electronic dispersion compensation for coherent optical channels," IEEE Transactions on Circuits Syst. I: Regul. Pap. **61**, 1012–1025 (2014).

32. I. Fatadin, D. Ives, and S. Savory, "Blind Equalization and Carrier Phase Recovery in a 16-QAM Optical Coherent System," J. Light. Technol. **27**, 3042–3049 (2009).

33. L. Schmalen, S. ten Brink, and A. Leven, "Advances in detection and error correction for coherent optical communications: Regular, irregular, and spatially coupled LDPC code designs," in *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks,* (John Wiley & Sons, Inc, 2016), pp. 65–122.

34. G. Liu, K. Zhang, R. Zhang, R. Proietti, H. Lu, and S. J. B. Yoo, "Demonstration of a carrier frequency offset estimator for 16-/32-QAM coherent receivers: a hardware perspective," Opt. Express **26**, 4853–4862 (2018).

35. M. Selmi, Y. Jaouen, and P. Ciblat, "Accurate digital frequency offset estimator for coherent polmux qam transmission systems," in *European Conference on Optical Communication (ECOC),* (2009), p. P3.08.

36. A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems," J. Light. Technol. **34**, 707–721 (2016).

37. B. Spinnler, S. Calabrò, and M. Kuschnerov, "Pilot-assisted channel estimation methods for coherent receivers," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013,* (Optical Society of America, 2013), p. OW4B.3.

38. D. Ashmawy, K. Banovic, E. Abdel-Raheem, M. Youssif, H. Mansour, and M. Mohanna, "Joint MCMA and DD blind equalization algorithm with variable-step size," in *2009 IEEE International Conference on Electro/Information Technology,* (2009), pp. 174–177.

39. P. M. Krummrich, D. Ronnenberg, W. Schairer, D. Wienold, F. Jenau, and M. Herrmann, "Demanding response time requirements on coherent receivers due to fast polarization rotations caused by lightning events," Opt. Express **24**, 12442–12457 (2016).

40. M. Karlsson, J. Brentel, and P. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," J. Light. Technol. **18**, 941–951 (2000).

41. K. Choutagunta, R. Ryf, N. Fontaine, S. Wittek, J. C. Alvarado-Zacarias, M. Mazur, H. Chen, R.-J. Essiambre, R. Amezcua-Correa, T. Hayashi, Y. Tamura, T. Hasegawa, T. Taru, and J. M. Kahn, "Modal dynamics in spatially multiplexed links," in *Optical Fiber Communication Conference (OFC) 2019,* (Optical Society of America, 2019), p. W4C.1.

42. A. Spalvieri and L. Barletta, "Pilot-aided carrier recovery in the presence of phase noise," IEEE Transactions on Commun. **59**, 1966–1974 (2011).

43. A. F. Alfredsson, R. Krishnan, and E. Agrell, "Joint-polarization phase-noise estimation and symbol detection for optical coherent receivers," J. Light. Technol. **34**, 4394–4405 (2016).

44. M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, M. Bertolini, and G. Gavioli, "Pilot-symbols-aided carrier-phase Recovery for 100-G PM-QPSK digital coherent receivers," IEEE Photonics Technol. Lett. **24**, 739–741 (2012).

45. E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," J. Light. Technol. **25**, 2675–2692 (2007).

46. D. S. Millar, R. Maher, D. Lavery, T. Koike-Akino, M. Pajovic, A. Alvarado, M. Paskov, K. Kojima, K. Parsons, B. C. Thomsen, S. J. Savory, and P. Bayvel, "Design of a 1 Tb/s superchannel coherent receiver," J. Light. Technol. **34**, 1453–1463 (2016).

47. T. Pfau, S. Hoffmann, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," J. Light. Technol. **27**, 989–999 (2009).

48. M. Mazur, A. Lorences-Riesgo, J. Schroder, P. A. Andrekson, and M. Karlsson, "High spectral efficiency PM-128QAM comb-based superchannel transmission enabled by a single shared optical pilot tone," J. Light. Technol. **36**, 1318–1325 (2018).

49. A. J. Metcalf, V. Torres-Company, D. E. Leaird, and A. M. Weiner, "High-power broadly tunable electrooptic frequency comb generator," J. Sel. Top. Quantum Electron. **19**, 231–236 (2013).

50. R. Dar, S. Chandrasekhar, A. H. Gnauck, B. Li, J. Cho, E. C. Burrows, and P. J. Winzer, "Impact of WDM channel correlations on nonlinear transmission," in *European Conference on Optical Communication (ECOC),* (2016), p. W.1.D.2.

51. J. Schröder and M. Mazur, "Chalmersphotonicslab/qampy: V0.1," (2018).

52. L. B. Mercer, "1/f frequency noise effects on self-heterodyne linewidth measurements," J. Light. Technol. **9**, 485–493 (1991).

53. R. Noe, "Phase noise-tolerant synchronous QPSK/BPSK baseband-type intradyne receiver concept with feedforward carrier recovery," J. Light. Technol. **23**, 802–808 (2005).

54. E. Börjeson, C. Fougstedt, and P. Larsson-Edefors, "ASIC design exploration of phase recovery algorithms for M-QAM fiber-optic systems," in *Optical Fiber Communication Conference (OFC) 2019,* (Optical Society of America, 2019), p. W3H.7.