Chalmers Publication Library

CHALMERS

Copyright Notice

©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This document was downloaded from Chalmers Publication Library (http://publications.lib.chalmers.se/), where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec. 8.1.9 (http://www.ieee.org/documents/opsmanual.pdf)

(Article begins on next page)

A Blind Phase Stabilization Algorithm for Parallel Coherent Receivers

Pontus Johannisson, Christophe Gosset, Magnus Karlsson

***Abstract*—The impact from phase drifts in the different branches of parallel coherent receivers is investigated and it is shown how the spectrum is broadened when the receiver branches are not phase stabilized. Based on this, we propose a blind algorithm for compensating these phase drifts in digital signal processing by minimization of the spectral width. The algorithm performance is then evaluated by numerical simulations of quadrature phase-shift keying data using return-to-zero modulation. It is found that the algorithm is capable of identifying the phases with sufficient accuracy to make the residual effect of the phase mismatches negligible compared to the signal distortion by noise at a bit error rate of 10^{-3}.**

***Index Terms*—Optical fiber communication, intradyne detection, digital signal processing.**

I. Introduction

COHERENT detection in combination with polarization multiplexing allows optical transmission with high spectral efficiency and makes it possible to use digital signal processing (DSP) to compensate for signal distortion. Although the development of the necessary receiver electronics is challenging, the DSP-based coherent receiver is a commercial reality since the demonstration of a 40 Gbit/s real-time system using dual-polarization quadrature phase-shift keying (QPSK) [1]. However, the analog-to-digital converter (ADC) is a bottleneck for the development of the next generation systems, operating at 400 Gbit/s–1 Tbit/s, since the bandwidth of even state-of-the-art ADCs is not sufficient for such high-speed optical signals. This is often a constraint already today when using optical time-division multiplexing or when the optical signal is generated using high-speed electrical signals [2], [3].

One way to work around the ADC bandwidth limitations is to perform optical time-division demultiplexing in the receiver by using a pulsed local oscillator (LO) [4]. This scheme works well and excellent results have been demonstrated both back-to-back and with transmission when the inter-symbol interference (ISI) is low, see for example [5], [6]. However, there are two drawbacks with this scheme. Firstly, the LO pulses must be synchronized with the signal symbols to obtain overlap with a specific tributary. Secondly, the scheme is limited to a single sample per symbol. This is because even if the bandwidth of the ADCs is compatible with the tributary symbol rate, it is typically not possible to resolve the symbol pulses, which are much shorter than the pulse spacing of a tributary. The sampled spectrum will therefore suffer from aliasing and it is impossible to compensate for arbitrary amounts of ISI such as chromatic dispersion (CD) and/or polarization-mode dispersion (PMD).

In order to introduce impairment mitigation capabilities in high symbol rate systems, it has been proposed to use a more general approach, which is here referred to as the *parallel coherent receiver* (PCR). With this type of receiver it is possible to sample at an arbitrary rate by using narrow LO pulses and sufficient parallelization to limit the ADC sampling rate in accordance with the ADC bandwidth. The history of this device goes back, at least, to the suggestion of using coherent detection with a pulsed LO as a way to do equivalent-time sampling of constellations with high bandwidth and sensitivity [7]. The next step was the introduction of a parallel structure for studying differential formats, although still by using a low LO pulse repetition frequency (10 MHz) [8]. This approach was then refined into a real-time capable configuration by raising the optical sampling frequency to two samples per symbol, thereby also enabling advanced DSP [9]. The PCR has then been further investigated and improved, which has lead to experimental demonstration of 4-fold parallelization and reception of QPSK data up to 64 Gbaud [10], [11]. This was achieved with an ADC bandwidth of 20 GHz, which shows the potential of the concept since reception and post-processing would not have been possible without parallelization.

Compared to the time-division demultiplexing approach, the sampling rate of the PCR can be set to be sufficient for adaptive equalization in DSP. Furthermore, it is not necessary to perform optical clock recovery to synchronize the signal to the LO pulses. The concept is also scalable in the sense that, in principle, it can be extended to any desired degree of parallelization, leading to arbitrarily low demands on the ADC bandwidth. This makes the PCR a candidate for receiver implementation in future transmission systems with high data rates, typically 400 Gbit/s–1 Tbit/s. Another promising application is in monitoring of high-speed optical signals. In this case, an increased bandwidth beyond what the conventional coherent receiver can achieve is highly valuable and an increased receiver cost can therefore be tolerated.

The PCR needs to be properly set up in order to achieve high performance. For example, the LO pulses should be narrow and have negligible chirp, and the ADCs should, ideally, be clocked synchronously with the LO pulses. This is beneficial since it allows the electrical sampling rate to be equal to

Pontus Johannisson and Magnus Karlsson are with the Photonics Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-412 96 Göteborg, Sweden, e-mail: pontus.johannisson@chalmers.se, magnus.karlsson@chalmers.se.

Christophe Gosset is with Telecom ParisTech, CNRS UMR LTCI, 46 rue Barrault, 75634 CEDEX Paris 13, France, e-mail: christophe.gosset@telecom-paristech.fr.

Manuscript received ??? ?, 201?; revised ??? ?, 201?.

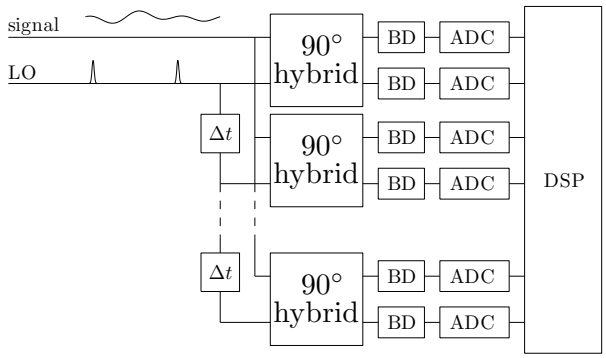

Fig. 1. Schematic picture of the PCR for a single-polarization case. BD = balanced detector. ADC = analog-to-digital converter. DSP = digital signal processing. For a dual-polarization signal, one more similar array of hybrids needs to be added.

the LO pulse repetition rate [11]. A practical complication in a non-integrated PCR (as compared to the conventional coherent receiver) is the phase shifts between the different receiver branches that are caused by slowly varying optical path lengths. However, provided that the phase differences can be identified, this can be compensated for in DSP, which would remove the demand for hardware phase stability. Such compensation is helpful in the system prototyping stage and can potentially even eliminate the need for integrated devices.

One algorithm for identifying the phase difference between the two receiver branches using 2-fold parallelization for the binary phase-shifting keying format was described in [9]. However, this algorithm is hard to generalize to higher degree of parallelization and other modulation formats. For recent experimental results [11], the algorithm from [12] was used. It is not clear exactly what objective function was used to run the optimization, but the authors explicitly state that there is a need for development of an integrated optical device or an improved algorithm to handle this problem [11].

In this paper we propose and investigate a blind algorithm for phase stabilization of a PCR. We start by demonstrating analytically how the spectrum of the sampled signal is affected by the type of phase shifts that occur in a PCR. Starting from this description, we suggest to estimate the phases by minimizing the spectral width. The performance of the proposed algorithm is then evaluated in numerical simulations and the results are discussed and concluded.

Notation: Vectors are denoted in bold letters (e.g., $\mathbf{u}$), and matrices in capital bold letters (e.g., $\mathbf{U}$). Element extraction from vectors or matrices are written with subscripts, e.g., $U_{m,n}$. Transpose is denoted by $\mathbf{a}^{\mathrm{T}}$, conjugation by $\mathbf{a}^{*}$, and the conjugate (Hermitian) transpose by $\mathbf{a}^{\mathrm{H}}$.

II. Receiver Description

A schematic picture of the PCR is seen in Fig. 1 and the operating principle is as follows: Assuming that the incoming signal is to be sampled at the rate f_{samp} and using N-fold parallelization, a pulsed LO is set up with repetition frequency f_{samp}/N. Sending the signal and the pulsed LO to a 90° hybrid we will, in effect, perform optical sampling before the ADC. By clocking the ADC synchronously with the LO pulse train, the electrical sampling rate can be set equal to the LO pulse repetition rate [11]. To obtain all samples, an array of N 90° hybrids must be used and the LO pulses are delayed $(n-1)/f_{\mathrm{samp}}$, $n \in [1, N]$ for the nth branch. The complete signal is obtained by interleaving the different sampled data streams.

As is easily worked out theoretically, using narrow LO pulses enables high-bandwidth optical sampling. The relative time delays in the different PCR branches must then be accurately tuned, and the clocking of the ADCs should be synchronous with the LO pulses. In a correctly configured system, the bandwidth demand on the ADCs is reduced by a factor N.

Using a non-integrated (fiber-based) PCR, it is impossible to avoid unequal phase shifts in the different receiver branches, i.e., during the propagation after having split the signal and the LO into multiple propagation paths. This means that neighboring samples in the interleaved sampled signal will have a random phase relation, which will cause conventional electronic dispersion compensation (EDC) or phase synchronization to fail. Therefore, the phase shifts need to be estimated at an early stage in the post-processing. Furthermore, since the lengths of the receiver branches change slowly (due, e.g., to temperature fluctuations), the phases drift with time. For continuous operation, this must be tracked but the drift time scale should be on the order of > 1 ms, i.e., much longer than the symbol slot. This relaxes the demands for the tracking speed and makes it possible to base the estimation on long data sequences.

III. Spectral Broadening from Phase Shifts

As a first step in designing the phase stabilization algorithm, we describe the impact on the signal spectrum from the type of phase shifts present in a PCR. Before describing the spectral broadening theoretically, we discuss a specific example.

Fig. 2a shows the power spectral density (PSD) of a 50% return-to-zero (RZ 50%) signal [13] with random QPSK data. By using a very high sampling rate, 16 samples per symbol in this example, the PSD is seen with very little aliasing and the impact on the spectrum is easy to see. The PSD of a signal sampled in one of the branches of a PCR using 4-fold parallelization is plotted in Fig. 2b. Due to the decrease in sampling rate, the frequency interval is here four times shorter than in Fig. 2a. In order to discuss the spectrum of the complete sampled signal, it is helpful to upsample the signal in one branch to f_{samp} by zero insertion. The PSD of the resulting signal is seen in Fig. 2c, and shows the periodic continuation of the PSD of Fig. 2b. The total signal is obtained by interleaving the signals sampled in the different branches. Since this is equivalent to adding four time-shifted upsampled signals, the spectrum of the interleaved signal is the superposition of four similar spectra corresponding to the PSD plotted in Fig. 2c. If the phase shifts of the PCR branches are all identical, then this results in the PSD from Fig. 2a. However, with non-ideal phase relation between the branches, the different spectra do not interfere in the right way. One example, calculated by using random phase shifts, is seen in

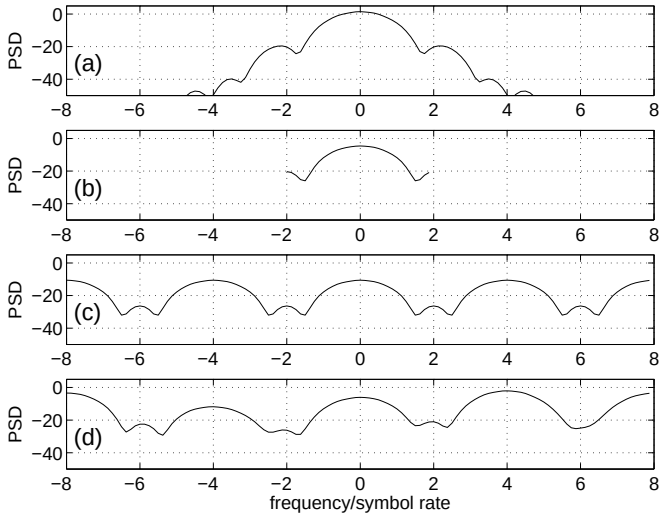

Fig. 2. (a) The PSD in dBm/(0.1 nm) for a RZ 50% QPSK signal. (b) The PSD of the signal sampled in one of the PCR branches. (c) The PSD of the signal from b upsampled by zero insertion. (d) The PSD of the interleaved signal with random phase shifts applied in the four receiver branches.

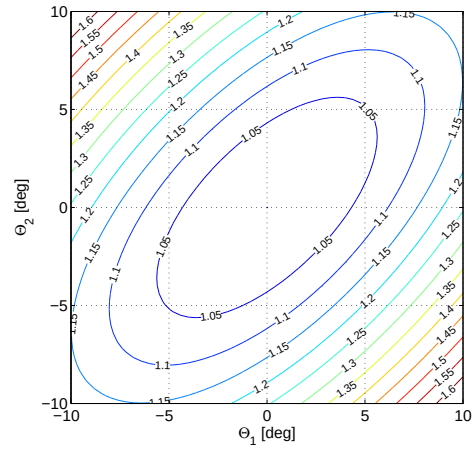

Fig. 3. A contour plot of the spectral width, corresponding to Fig. 2, as a function of the phase deviation measured in degrees from the correct value. The width has been normalized to its minimum value. In this case $\Theta_3 = \Theta_4 = 0$. The minimum is very close to $\Theta_1 = \Theta_2 = 0$.

Fig. 2d. The result is qualitatively somewhat similar to Fig. 2c, but far from the correct PSD of Fig. 2a.

To describe the impact on the spectrum theoretically, we denote the $(M \times 1)$ interleaved sampled signal in the absence of any phase shifts, i.e., sampled by an ideal PCR, by $\mathbf{u}$. Introducing a phase shift Θ_n in the nth branch, we set up a phase shift vector $\boldsymbol{\psi} = [e^{i\Theta_1}, e^{i\Theta_2}, \ldots, e^{i\Theta_N}]^{\mathrm{T}}$. To a good approximation, $\boldsymbol{\psi}$ is constant over the time we consider. Allowing ψ^{e}_m, $m = 1, \ldots, M$ to be a periodic continuation of $\boldsymbol{\psi}$, we can express the sampled signal in the presence of the phase shifts as $\tilde{u}_m = \psi^{\mathrm{e}}_m u_m$. Since $\boldsymbol{\psi}^{\mathrm{e}}$ has the period N, it can be written as

$$\psi^{\mathrm{e}}_m = \sum_{n=1}^{N} c_n e^{i\Omega_n (m-1)}, \tag{1}$$

where $\Omega_n = 2\pi(n-1)/N$ is the angular frequency of the nth Fourier component and the (unknown) Fourier coefficients c_n depend only on the phase shifts Θ_n. We notice that if the phase shift is equal in all the PCR branches, then $|c_1| = 1$ and $c_n = 0$, $n > 1$. However, in the general case, $c_n \neq 0$, $\forall n$, such that

$$\tilde{u}_m = u_m \sum_{n=1}^{N} c_n e^{i\Omega_n (m-1)}. \tag{2}$$

From this expression, we can identify the impact on the spectrum from the phase shifts in a PCR: Performing the discrete Fourier transform (DFT) of $\tilde{\mathbf{u}}$ from (2), we will find a superposition of scaled (by c_n), frequency translated (by Ω_n) replicas of the spectrum of $\mathbf{u}$. The true spectrum will be recovered only when the phase shift is equal in all branches.

From (2) and Fig. 2d we understand that EDC will not work as intended in the presence of phase shifts, since only the part of the spectrum that is still centered at zero frequency will be correctly compensated by the parabolic phase shift introduced by EDC. Phase compensation should therefore be done before EDC is carried out. Anticipating the result of the next section, we emphasize that CD only affects the spectral phase, not the amplitude.

IV. The Phase Stabilization Algorithm

The spectral shape of optical signals used in coherent communication is well known for many digital modulation formats [13], [14]. One important example is the spectrum of a return-to-zero quadrature amplitude modulation (QAM) signal, which has been illustrated for the specific choice of QPSK in Fig. 2a. As is intuitively clear from Fig. 2d, the phase shifts in the different PCR branches lead to a spectral broadening. We therefore suggest that for many modulation formats of practical importance in coherent optical communication, the phase stabilization can be done by minimizing the spectral width. For a continuous signal, the root mean square (RMS) spectral width is defined from

$$\Delta f^2_{\mathrm{RMS}} = \int_{-\infty}^{\infty} f^2 |X(f)|^2 \, df \Big/ \int_{-\infty}^{\infty} |X(f)|^2 \, df, \tag{3}$$

where $X(f)$ is the Fourier transform of an arbitrary (transformable) signal $x(t)$.

In order to show that the approach is promising, the spectral width as a function of the PCR branch phases is seen as a contour plot in Fig. 3. Without loss of generality, the fourth phase is set to zero. In this specific case also $\Theta_3 = 0$ due to the difficulty to visualize a three-dimensional function. However, the result is qualitatively identical for other choices. The drawn out shape of the contours occurs since the width increases more slowly when $\Theta_1 = \Theta_2$. By increasing the span of the deviation phases to $[-\pi, \pi)$, it has been found that there is a unique minimum.

If the spectrum is very flat, the suggested approach can fail. Unfortunately, we have not been able to find a strict condition when this will be the case. It is, for example, not sufficient that the amplitude of the spectrum is monotonically decreasing with the magnitude of the frequency. An example of this in

the case $N = M = 2$ is the spectrum vector $[1, 0.9i]^{\mathrm{T}}$. The algorithm will in this case be able to shift almost all power to DC. However, this example is very far from being realistic. In Section VI, we have investigated the performance numerically using realistic parameters and the approach is then proven by the fact that the result is very good.

A. The Minimization Problem

The proposed spectral width minimization will be based on the analysis of the spectrum of the sampled signal $\tilde{\mathbf{u}}$. The aim is that the corrected version, $\hat{\mathbf{u}}$, of $\tilde{\mathbf{u}}$ should be as similar as possible to the signal $\mathbf{u}$, which would be obtained by sampling with an ideal PCR. The corrected signal has the corresponding DFT $\hat{\mathbf{v}}$, and we write the Fourier transformation operation as the linear transform $\hat{\mathbf{v}} = \mathbf{W}\hat{\mathbf{u}}$, where the $(M \times M)$ matrix $\mathbf{W}$ is given by the definition of the DFT. A discrete counterpart of (3) is given by (the square of) the normalized RMS width of the DFT according to

$$\Delta f_{\mathrm{RMS}}^2 = \hat{\mathbf{v}}^{\mathrm{H}}\mathbf{G}\hat{\mathbf{v}}/(\hat{\mathbf{v}}^{\mathrm{H}}\hat{\mathbf{v}}), \tag{4}$$

where the diagonal matrix $\mathbf{G}$ contains the elements $G_{m,m} = f_m^2$, and f_m is the normalized frequency in the interval $[-1, (M-1)/M]$ corresponding to the DFT element $\hat{v}_m$. Since $\hat{\mathbf{v}}^{\mathrm{H}}\hat{\mathbf{v}}$ is not affected by the phase shift, we can choose to define a cost function for spectral minimization according to

$$J = \hat{\mathbf{v}}^{\mathrm{H}}\mathbf{G}\hat{\mathbf{v}}. \tag{5}$$

We introduce the $(M \times N)$ matrix $\tilde{\mathbf{U}}$ with elements $\tilde{U}_{m,n} = P_{m,n}\tilde{u}_m$, where

$$P_{m,n} = \begin{cases} 1, & m = n + aN \text{ for some integer } a, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

In this way, the columns of $\tilde{\mathbf{U}}$ contain the signals from the different branches upsampled to f_{samp} by zero insertion. As opposed to Section III, we now introduce deliberate phase compensation in each branch and denote the phase shift vector by $\boldsymbol{\phi} = [e^{i\theta_1}, e^{i\theta_2}, \ldots, e^{i\theta_N}]^{\mathrm{T}}$ (compare with $\boldsymbol{\psi}$ from Section III). The corrected vector can then be written $\hat{\mathbf{u}} = \tilde{\mathbf{U}}\boldsymbol{\phi}$. Using that

$$J = (\mathbf{W}\tilde{\mathbf{U}}\boldsymbol{\phi})^{\mathrm{H}}\mathbf{G}(\mathbf{W}\tilde{\mathbf{U}}\boldsymbol{\phi}), \tag{7}$$

we can formulate the optimization problem as

$$\hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}}(\boldsymbol{\phi}^{\mathrm{H}}\mathbf{A}\boldsymbol{\phi}), \tag{8}$$

where

$$\mathbf{A} = \tilde{\mathbf{U}}^{\mathrm{H}}\mathbf{W}^{\mathrm{H}}\mathbf{G}\mathbf{W}\tilde{\mathbf{U}}. \tag{9}$$

Here, $\hat{\boldsymbol{\phi}}$ is the optimal phase compensation vector that minimizes the cost function J, and thus also the spectral width (4), and the $(N \times N)$ Hermitian matrix $\mathbf{A}$ is calculated from the sampled data. It is worth noticing that the columns of $\mathbf{W}\tilde{\mathbf{U}}$ are the DFTs of the upsampled signals of the individual branches. The upsampling by zero insertion will simply cause an N-fold periodic continuation of the spectra, as was illustrated in Fig. 2c. We point out that although many samples are needed to set up the $\mathbf{A}$ matrix, the resulting matrix is small, $(N \times N)$, and the following calculation of the $\hat{\boldsymbol{\phi}}$ vector does not involve any further processing of the sampled data.

Before continuing, we remark that (9) takes the entire spectrum into consideration. One way to avoid the DFT operations would be to only consider the spectral components centered at Ω_l (defined in the discussion of (1)). These could be extracted by a bank of bandpass filters and would simplify the setup of $\mathbf{A}$.

B. Solution by Steepest Descent

The phase compensation vector $\hat{\boldsymbol{\phi}}$ is obtained from (8). Although it may be possible to find the solution with analytical methods, we proceed in a simple way by using the steepest descent method. Since the elements of $\hat{\boldsymbol{\phi}}$ should have normalized modulus, we extract the phases into the vector $\hat{\boldsymbol{\theta}} = [\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N]^{\mathrm{T}}$, which has no such constraints. We then update $\hat{\boldsymbol{\theta}}$ from iteration k to $k+1$ by taking a step in the opposite direction of the gradient according to

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k - \mu\nabla_{\hat{\boldsymbol{\theta}}} J(\hat{\boldsymbol{\theta}}_k), \tag{10}$$

where μ is the step size, and $\nabla_{\hat{\boldsymbol{\theta}}}$ is the gradient obtained by differentiating the cost function with respect to the vector $\hat{\boldsymbol{\theta}}$, with the result

$$\nabla_{\hat{\boldsymbol{\theta}}} J = 2\,\mathrm{Im}[\mathrm{diag}(\mathbf{A}\hat{\boldsymbol{\phi}}\hat{\boldsymbol{\phi}}^{\mathrm{H}})], \tag{11}$$

where diag extracts the diagonal elements into a column vector. This allows the vector $\hat{\boldsymbol{\phi}}$ to be estimated and then tracked by using updated $\mathbf{A}$ matrices.

V. System Aspects

It is an important fact that the suggested algorithm can be included into the typical chain of DSP in the receiver. To allow normal operation of the subsequent DSP, the phase shifts should be compensated early, preferably just after compensation of branch-specific hardware errors. This is possible since the phase stabilization can be performed on asynchronously sampled data and the algorithm is not affected by CD or PMD. The latter is because only the amplitude of the spectrum is used, but not the phase. Polarization-dependent loss will also have no impact as long as it has no frequency dependence.

Furthermore, it is important that the signal after compensation is such that the subsequent DSP will be able to work as normal. It is a very difficult question how this will work in detail since the answer depends on the algorithm choice. For example, the constant modulus algorithm will not be affected by the phase distortion since its cost function only depends on the signal amplitude. On the other hand, the phase synchronization algorithm will obviously be affected, but the impact may vary depending on the algorithm. Here, we therefore instead start from the approach that the signal distortion due to residual phase shifts should be small compared to the noise typically present in the signal. When this is true, we expect the impact on the subsequent DSP to be very small, but the detailed analysis of this is outside the scope of this work.

In order to be able to compare error sources we use the error vector magnitude (EVM) defined according to

$$\mathrm{EVM} = \sqrt{\mathbb{E}\,|\hat{u} - u|^2}, \tag{12}$$

where u is the ideal signal, $\hat{u}$ is the distorted signal, and the expectation operator is denoted by $\mathbb{E}$. In the case of a noisy but otherwise ideal signal, we will directly find that $\hat{u}-u=n$, where n is the noise, and we get

$$\mathrm{EVM}_{\mathrm{noise}} = \sqrt{P_{\mathrm{noise}}}, \tag{13}$$

where P_{noise} is the average noise power. Instead looking at an ideal signal with residual phase distortion, we have that

$$\hat{u}_k - u_k = u_k e^{i\Delta\theta_\kappa} - u_k, \tag{14}$$

where $\Delta\theta_\kappa$ is the difference of the exact and estimated phases and $\kappa \in [1, N]$ is obtained from $\kappa = k + aN$ for a suitably chosen integer a. Inserting this into (12), we obtain

$$\begin{aligned}\mathrm{EVM}_{\mathrm{phase}} &= \sqrt{P_{\mathrm{signal}}}\sqrt{\mathbb{E}\,|e^{i\,\Delta\theta_\kappa}-1|^2} \\ &\approx \sqrt{P_{\mathrm{signal}}}\sqrt{\mathbb{E}\,|\Delta\theta_\kappa|^2} = \sqrt{P_{\mathrm{signal}}}\,\sigma_{\Delta\theta},\end{aligned} \tag{15}$$

where $\sigma_{\Delta\theta}$ is the standard deviation of $\Delta\theta$ and we have used that $\Delta\theta$ is a small angle. The ratio of these two amplitude errors is

$$\eta = \frac{\mathrm{EVM}_{\mathrm{phase}}}{\mathrm{EVM}_{\mathrm{noise}}} = \sigma_{\Delta\theta}\sqrt{\frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}}} = \sigma_{\Delta\theta}\sqrt{\mathrm{SNR}}. \tag{16}$$

If we set the signal-to-noise ratio (SNR) 9.8 dB, which is required to obtain a bit error rate (BER) of 10^{-3} for a QPSK signal, and a standard deviation $\sigma_{\Delta\theta} = 1^\circ$, we obtain $\eta \approx 0.054$. Furthermore, $\eta < 0.1$ up to SNR levels that correspond to BER $< 10^{-8}$. If we instead consider rectangular 16-QAM, the SNR requirement at BER $= 10^{-3}$ increases by roughly 6 dB. Due to the square root in (16), we then obtain $\eta \approx 0.11$ This shows that a standard deviation of 1° is a reasonable accuracy target for practically relevant signals.

VI. Numerical Simulations

The performance of the proposed algorithm has been tested in numerical simulations. We first report the results from simulations where the system has negligible nonlinearities and the balanced detectors (BDs)/ADCs have sufficient bandwidth to cause negligible ISI (which here rather should be interpreted as inter-*sample* interference due to the parallelization.). These assumptions are then discussed in Section VII.

We find it convenient to investigate return-to-zero (RZ) optical signals, since they have a mathematically well-defined description and therefore do not rely on a specific choice of pulse-shaping filter [13]. By simulating different duty cycles, we examine how the required number of samples per symbol is affected by the spectral width. Random QPSK data has been used in all cases. The results are valid for arbitrary symbol rates since this will only cause a rescaling of the spectral width.

Since the system is linear, the signal quality is fully characterized by the SNR. Furthermore, since the result is not affected by CD/PMD, the system length does not need to be selected. The laser phase noise has been neglected since it has minimal impact on the spectrum amplitude. The reason is that the linewidth is of the order ~ 1 MHz, while the signal bandwidth is likely to be significantly larger than 10 GHz. Thus, the spectral broadening due to this effect is negligible.

Since the number of samples used to calculate the matrix $\mathbf{A}$ affects the result, we have tested 2^{12}, 2^{14}, and 2^{16} samples. The result also depends on the SNR and we have tested two different cases: one case with negligible noise and one case where complex additive white Gaussian noise (AWGN) has been added to make the mean noise power one tenth of the signal power. (The required SNR at BER $= 10^{-3}$ is 9.8 dB for QPSK.)

We have investigated the two cases of 4-fold and 8-fold parallelization. In the 4-fold case, the initial phases were all set to zero and then the algorithm was iterated until convergence was seen. The step size was chosen to be $\mu = 5/(\tilde{\mathbf{v}}^{\mathrm{H}}\tilde{\mathbf{v}})$, where $\tilde{\mathbf{v}}$ is the DFT of the sampled signal $\tilde{\mathbf{u}}$. The necessary number of iterations varies depending on, e.g., the number of samples per symbol. Typically, when the final standard deviation is roughly 1°, then well under 100 iterations are necessary. As opposed to the 4-fold case, the algorithm tends to find local minima in the 8-fold case. This presents a problem when starting up the algorithm but will not be a problem in the subsequent tracking of the phase evolution. We have chosen to work around this by evaluating the cost function for a number of candidate $\boldsymbol{\phi}$ vectors, and then proceed the iteration using the candidate with the lowest cost. By testing four individually chosen and equally spaced angles for each branch, we have been able to eliminate the problem. Typically, a couple of hundred iterations are needed in the 8-fold case.

To quantify the result, we have calculated the standard deviation, $\sigma_{\Delta\theta}$, of the phase error, i.e., the difference of the true and estimated phases, averaged over all receiver branches in 1000 simulations. The receiver branch phases have, in every simulation, been chosen to be independent and uniformly random in the interval $[-\pi, \pi)$.

The results for 4-fold parallelization are seen in Fig. 4 and the corresponding results using 8-fold parallelization are seen in Fig. 5. The three subfigures show, from the top, the results from using RZ 67%, RZ 50%, and RZ 33%, respectively. The thick lines indicate the result with noise included and the different markers indicate the amount of data that has gone into the calculation of the $\mathbf{A}$ matrix, as indicated in the legend. We have used asynchronous sampling and the number of samples per symbol is an irrational number. The axes has been chosen to show the result in the interval 0.1°–10°.

VII. Discussion

The results show that a standard deviation $\sigma_{\Delta\theta} = 1^\circ$ can be reached and that sufficient sampling rate is important for good performance. In particular, this is true for the modulation format with largest spectral width (RZ 33%) and is a natural consequence from the fact that a low sampling rate leads to aliasing and a flatter spectrum. As expected, the introduction of noise leads to increased phase standard deviation, but the impact is rather limited. In all cases, it is possible to increase the performance by using more data in the computation of the $\mathbf{A}$ matrix. However, for a low sampling rate, the necessary amount of data can become unpractically large.

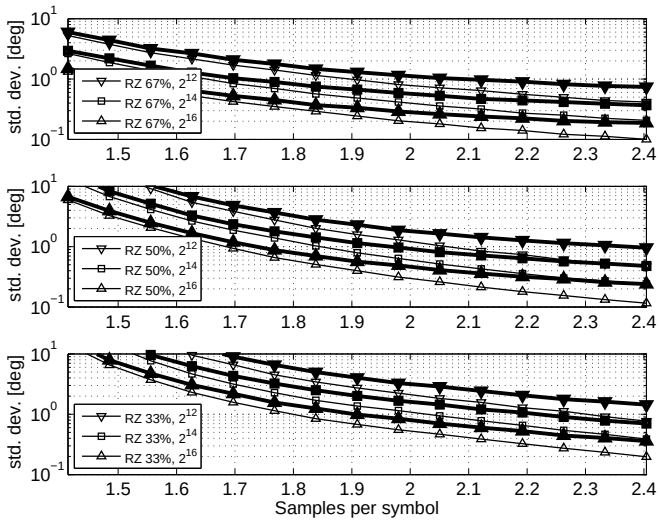

Fig. 4. The resulting phase standard deviation with 4-fold parallelization. Counted from the top, the figures show RZ 67%, RZ 50%, and RZ 33% modulation, respectively. The thin lines show the noiseless case. The markers (see legend) indicate the number of samples used to calculate the $\mathbf{A}$ matrix.

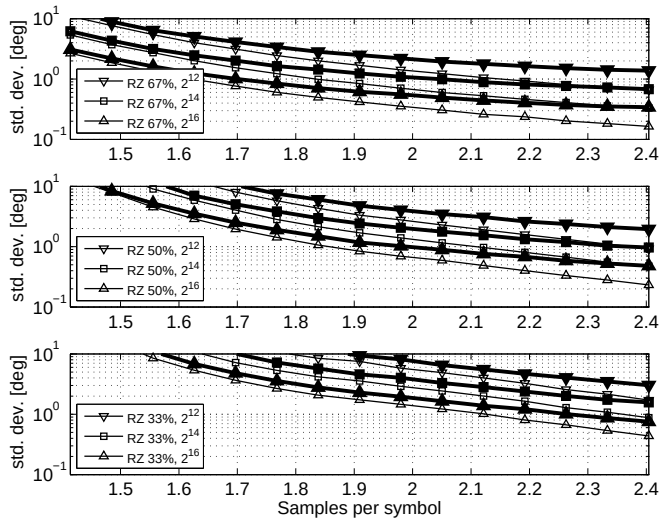

Fig. 5. The resulting phase standard deviation with 8-fold parallelization. Counted from the top, the figures show RZ 67%, RZ 50%, and RZ 33% modulation, respectively. The thin lines show the noiseless case. The markers (see legend) indicate the number of samples used to calculate the $\mathbf{A}$ matrix.

The results are qualitatively similar in all cases, also when comparing the 4-fold and the 8-fold cases. In general, the standard deviation is higher for the 8-fold case, which is explained by the fact that in this case we have more unknown parameters but we use the same amount of data. Since the signal absolute phase is of no consequence, there are 3 unknown parameters in the first case and 7 in the second case. As an example, at 2.0 samples per symbol, the simulations in the 8-fold case with noise included show a standard deviation increase of 1.9–2.3 times compared to the 4-fold case. However, using sufficient amount of data, accurate estimation of the phases is possible also with high degree of parallelization.

In order to show the distribution of the residual phase errors, Fig. 6 shows the numerically obtained histogram for the selected case of 4-fold parallelization, RZ 50%, 2^{14} samples obtained at 2 samples per symbol with noise included. In this case, the standard deviation is $0.96°$, and a fitted normal distribution has been included for comparison. It is seen that the probability density function of the residual phase error is Gaussian, which means that the standard deviation reported in the numerical simulations contains sufficient statistical information.

In the numerical simulations, it was assumed that the BDs/ADCs have sufficient bandwidth to cause negligible ISI. If this is not the case, the frequency components around ± 2 in our example Fig. 2b will be suppressed. In the interleaved signal, corresponding to Fig. 2a, the effect of the limited bandwidth will be seen around ± 2 and ± 6. The impact can be important at lower sampling rate, say 2 samples per symbol, since N dips will appear in the spectrum. This is a general feature of the PCR and will not be discussed in detail here. However, the impact on the algorithm performance is important to investigate. By lowpass filtering the signals in the different receiver branches, we have found that the algorithm is not very sensitive to this type of filtering. As an example, we have used a 5th order Butterworth filter with the cut-off frequency selected as to allow only half of the spectrum to pass. For the parameters used to produce Fig. 6, this increased the standard deviation from $0.96°$ to $1.35°$. A similar effect, i.e., around 40% increase, was also observed for the other modulation formats. The spectrum is in this case strongly distorted with 4 wide dips appearing. This suppression of signal content decreases the information available to the algorithm and causes the observed increase of the residual phase errors.

Above, the system was assumed to be linear but as is well known, the Kerr nonlinearity is an important system limitation in practice. Unfortunately, it is very hard to describe the impact of the Kerr nonlinearity generally and a detailed study of how this will affect the presented algorithm is outside of the scope of the paper. However, the PCR will most likely be used at very high symbol rates. It has been shown that in uncompensated WDM transmission, the Kerr nonlinearity will cause signal distortion that is statistically similar to AWGN [15]. This approach has led to a very good model for WDM transmission [16], and this means that the Kerr nonlinearity will simply reduce the "effective SNR". For this case, the above simulations are still usable. As a complementary case we consider a single-channel system with densely spaced optical dispersion compensation. In such a system, self-phase modulation is the most important nonlinear distortion. This leads to a broadening of the spectrum and also reduces the curvature, which leads to larger residual phase errors. Again, we have found that the impact is weak. In the case where the maximum nonlinear phase shift is 1 rad we found a 27% increase of the standard deviation using the parameters used to produce Fig. 6. Similar results were obtained for the other modulation formats.

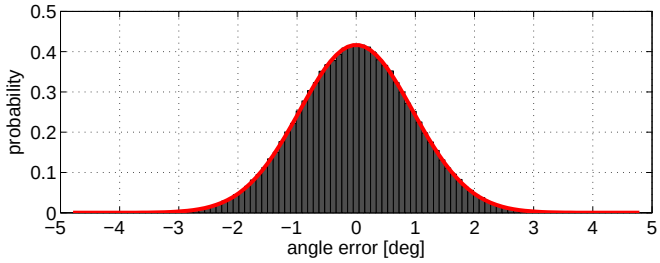

Fig. 6. The histogram shows the numerically obtained distribution of the residual phase errors for 4-fold parallelization, RZ 50%, 2^{14} samples obtained at 2 samples per symbol with noise included. The red curve shows a fitted normal distribution for comparison.

VIII. Conclusion

We have proposed an algorithm for phase stabilization of the parallel coherent receiver based on minimization of the spectral width. Numerical simulations have shown that the phase standard deviation after convergence is small and that the impact of the residual phase errors is negligible in comparison to the noise in many signals of practical importance.

Acknowledgment

We acknowledge the financial support from the EU EUROFOS project, the Swedish Research Council, and the Vinnova IKT program.

References

[1] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Opt. Express*, vol. 16, no. 2, pp. 873–879, Jan. 2008.

[2] P. J. Winzer, A. H. Gnauck, S. Chandrasekhar, S. Draving, J. Evangelista, and B. Zhu, "Generation and 1200-km transmission of 448-Gb/s ETDM 56-Gbaud PDM 16-QAM using a single I/Q modulator," in *European Conference on Optical Communication (ECOC)*, 2010, p. PD2.2.

[3] G.-W. Lu, M. Sköld, P. Johannisson, J. Zhao, M. Sjödin, H. Sunnerud, M. Westlund, A. Ellis, and P. A. Andrekson, "40-Gbaud 16-QAM transmitter using tandem IQ modulators with binary driving electronic signals," *Opt. Express*, vol. 18, no. 22, pp. 23 062–23 069, Oct. 2010.

[4] C. Zhang, Y. Mori, K. Igarashi, K. Katoh, and K. Kikuchi, "Ultrafast operation of digital coherent receivers using their time-division demultiplexing function," *J. Lightw. Technol.*, vol. 27, no. 3, pp. 224–232, Feb. 2009.

[5] ——, "Demodulation of 1.28-Tbit/s polarization-multiplexed 16-QAM signals on a single carrier with digital coherent receiver," in *Optical Fiber Communication Conference (OFC)*, 2009, p. OTuG3.

[6] C. Schmidt-Langhorst, R. Ludwig, D.-D. Groß, L. Molle, M. Seimetz, R. Freund, and C. Schubert, "Generation and coherent time-division demultiplexing of up to 5.1 Tb/s single-channel 8-PSK and 16-QAM signals," in *Optical Fiber Communication Conference (OFC)*, 2009, p. PDPC6.

[7] C. Dorrer, D. C. Kilper, H. R. Stuart, G. Raybon, and M. G. Raymer, "Linear optical sampling," *IEEE Photon. Technol. Lett.*, vol. 15, no. 12, pp. 1746–1748, Dec. 2003.

[8] K. Okamoto and F. Ito, "Ultrafast measurement of optical DPSK signals using 1-symbol delayed dual-channel linear optical sampling," *IEEE Photon. Technol. Lett.*, vol. 20, no. 11, pp. 948–950, June 2008.

[9] X. Chen, X. Xie, I. Kim, G. Li, H. Zhang, and B. Zhou, "Coherent detection using optical time-domain sampling," *IEEE Photon. Technol. Lett.*, vol. 21, no. 5, pp. 286–288, Mar. 2009.

[10] J. K. Fischer, R. Ludwig, L. Molle, C. Schmidt-Langhorst, A. Galperin, T. Richter, C. C. Leonhardt, A. Matiss, and C. Schubert, "High-speed digital coherent receiver with parallel optical sampling," in *Optical Fiber Communication Conference (OFC)*, 2010, p. PDPB4.

[11] J. K. Fischer, R. Ludwig, L. Molle, C. Schmidt-Langhorst, C. C. Leonhardt, A. Matiss, and C. Schubert, "High-speed digital coherent receiver based on parallel optical sampling," *J. Lightw. Technol.*, vol. 29, no. 4, pp. 378–385, Feb. 2011.

[12] L. D. Coelho, O. Gaete, and N. Hanik, "An algorithm for global optimization of optical communication systems," *Int. J. Electron. Commun.*, vol. 63, no. 7, pp. 541–550, July 2009.

[13] E. Ip and J. M. Kahn, "Power spectra of return-to-zero optical signals," *J. Lightw. Technol.*, vol. 24, no. 3, pp. 1610–1617, Mar. 2006.

[14] K.-P. Ho and J. M. Kahn, "Spectrum of externally modulated optical signals," *J. Lightw. Technol.*, vol. 22, no. 2, pp. 658–663, Feb. 2004.

[15] A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. Tapia Taiba, and F. Forghieri, "Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links," in *European Conference on Optical Communication (ECOC)*, 2010, p. P4.07.

[16] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, June 2011.

Pontus Johannisson received his Ph.D. degree from Chalmers University of Technology, Göteborg, Sweden, in 2006. His thesis was focused on nonlinear intrachannel signal impairments in optical fiber communications systems. In 2006, he joined the research institute IMEGO in Göteborg, Sweden, where he worked with digital signal processing for inertial navigation with MEMS-based accelerometers and gyroscopes. In 2009, he joined the Photonics Laboratory, Chalmers University of Technology, where he is currently Assistant Professor. His research interests include nonlinear effects in optical fibers and digital signal processing in coherent optical receivers.

Christophe Gosset received the Ph.D. degree in applied physics from the Ecole Nationale Supérieure des Télécommunications (ENST), presently Telecom ParisTech (Paris, France), in 2002. The subject of the thesis was four wave mixing in semiconductor optical amplifiers and lasers. From 2004 to 2006, he was a Postdoctoral researcher at the Laboratory for Photonic and Nanostructures - CNRS (Marcoussis, France), on the study of passive mode-locking in quantum-wells and quantum-dots semiconductor lasers, and on the pulse characterization at very high repetition rate. From 2007 to 2008 he was with Orange Labs (Lannion, France), in the Core Network Department with a research topic dedicated to 40–100 Gbit/s systems. Since 2008, he is a researcher at Telecom ParisTech. His research activities lie on advanced modulation formats and digital signal processing for high symbol rate optical transmissions, and high speed optical signal characterization.

Magnus Karlsson received his Ph.D. in Electromagnetic Field Theory in 1994 from Chalmers University of Technology, Göteborg, Sweden. The title of his Ph.D. thesis was "Nonlinear propagation of optical pulses and beams". Since 1995, he has been with the Photonics Laboratory at Chalmers, first as Assistant Professor and since 2003 as Professor in photonics. He has authored or co-authored over 190 scientific journal and conference contributions, served as guest editor for the Journal of Lightwave Technology, and is currently associate editor of Optics Express. He has served in the technical committees for the Optical Fiber Communication Conference (OFC) (2009 as subcommittee chair), and the Asia Communications and Photonics Conference (ACP, formerly APOC).

His research has been devoted to a variety of aspects of fiber optic communication systems, in particular transmission effects such as fiber nonlinearities and polarization effects, but also applied issues such as high-capacity data transmission and all-optical switching. Currently he is devoted to parametric amplification, multilevel modulation formats, and coherent transmission in optical fibers.