

## System of algorithms for digital signal processing for coherent optical communications

© T.O. Bazarov,<sup>1</sup> M.A. Senko,<sup>1,2</sup> L.A. Samodelkin,<sup>1,2</sup> I.S. Khalko,<sup>1,3</sup> I.N. Solomadin,<sup>1</sup> O.E. Naniy,<sup>1,2</sup> V.N. Treschikov<sup>1</sup>

<sup>1</sup> T8 Company, Moscow, Russia

<sup>2</sup> Moscow State University, Moscow, Russia

<sup>3</sup> Moscow Institute of Physics and Technology, Dolgoprudny, Moscow oblast, Russia

e-mail: timur.bazarov@phystech.edu

Received January 10, 2024 Revised March 15, 2024 Accepted March 21, 2024

One of the main parts of modern coherent fiber-optic transceivers is the digital signal processing unit (DSP). Its task is to compensate for the distortions of the transmitted signal occurring at the transmitter, in the fiber-optic communication line (FOCL) and at the receiver. This paper presents an analysis of existing linear DSP algorithms that allow to recover the DP-QPSK (dual polarization quadrature phase shift–keying) signal. The original implementation of the complete digital signal processing path of the receiving part is also demonstrated, tested on experimental signal data, which passed the line up to 4000 km. Demodulation results are compared with the theoretical limit and the best available commercial device.

**Keywords:** Digital signal processing, equalization, coherent communications, digital filtration, DP-QPSK.

DOI: 10.61011/TP.2024.06.58826.3-24

### Introduction

Some of the main current trends are widespread digitization and big data analysis, which causes the fastest growth of transmitted information volume in history. According to data of CISCO [1] for the period from 2015 to 2020 the global IP traffic tripped from 72.5 to 194.3 exabytes ( $10^{18}$ ) per month. The growth in the number of Internet users and average data transmission speed also continues [2]. The main technology that makes it possible to meet the permanent demand for the expansion of the throughput capacity of the communication lines is a fiber coherent optics.

Signal transmission along optic fiber is a dominant technology in the market and occupies around 90% of all wideband communication lines [3]. Such share became possible because of introduction of some technologies: coherent detection demonstrated in the 80s of the 21st century [4], DWDM — dense wavelength division multiplexing, and EDFA — erbium doped fiber amplifier [5].

Initially in fiber optic lines (FOLs) the amplitude modulation of laser radiation was used for information coding, and direct detection was used for reception. Channel speed in the commercial systems with direct modulation reached 10 Gbit/s (STM-64) in 1995 and 40 Gbit/s (STM-256) in 2002. However, the development of systems with the direct modulation did not stop there, since the band width of the amplitude-modulated signal approached the width of the available spectral band in the frequency grid ITU-T.

It was necessary to increase the information capacity of the signal, which was implemented by introduction of amplitude-phase coding and coherent detection. The principle of coherent reception consists in the fact that the

signal received along the line is mixed with the radiation of the reference laser at nearby frequency. The differential signal contains information both about amplitude and phase of the information signal. However, such reception provides not only the additional degrees of freedom, but increases requirements to the quality of signal distortion compensation. Some of such distortions are phase distortions caused by the end width of the laser radiation line on the receiver and the transmitter, and their frequency offset from each other. Such distortions may be excluded, if one and the same laser is used as a reference laser for transmission and reception [6]. However, in such a case the laser signal should pass the line close in length to the one that is passed by the information signal, which is bad both from the point of view of increasing the necessary number of fibers, and from the point of view of the transmission distance. Therefore, prior to demonstration of the digital methods for compensation of the phase distortions of the laser [7], the coherent optical communication has not practically developed.

Use of the digital signal processing (DSP) for compensation of distortions accumulated in the signal made it possible to both use the coherent detection in practice [7], and to use two orthogonal polarizations to transmit information, to introduce complicated formats of modulation [8–10] and to considerably increase both the volume of the transmitted information and the transmission distance.

Intense research of coherent detection methods were carried out in the 90s with the purpose to implement their potential benefits in the sensitivity and speed of information transmission, real-time commercial solutions appeared in 2000s [11].

Signal transmission along FOLs is accompanied with linear and non-linear distortions of the signal arising to a

certain extent in all components of FOLs. The main linear distortions arising in the line are the following: chromatic dispersion, polarization mode dispersion (PMD), polarization dependent losses, quick polarization rotation, noise of amplified spontaneous emission (ASE), optical filtration, unevenness of attenuation and gain spectra. Besides, there are also distortions caused by a transmitter and a receiver, such as timing mal-synchronization, frequency offset, phase noises of the laser, and their amplitude-frequency response, phase frequency response and quadrature-to-quadrature delays. DSP is aimed at fighting against all the above distortions, except for ASE noise, which in the channel band is white and is summed up with the signal additively. The main method to fight against ASE is using error control coding (FEC — forward error correction). At error control coding with redundancy of around 15–20% in the current coherent receivers it becomes possible to reduce the share of erroneous bits (BER — bit error ratio) from level  $\sim 10^{-2}$  upstream FEC block to  $10^{-15}$  downstream FEC block [12].

For the linear systems the objective of DSP may be formalized as follows: let  $\{a_n\}$  — transmitted data;  $F_s = 1/T_s$  — sampling rate of an analog-to-digital converter (ADC);  $h(kT_s)$  — full impulse response of the system, including a shaping filter, a response of a transmitter, a line and a receiver after an ADC;  $v(kT_s)$  — sampled noise of ASE; then the signal after ADC appears as

$$y_k = \sum_n a_n h(kT_s - nT) + v(kT_s), \quad (1)$$

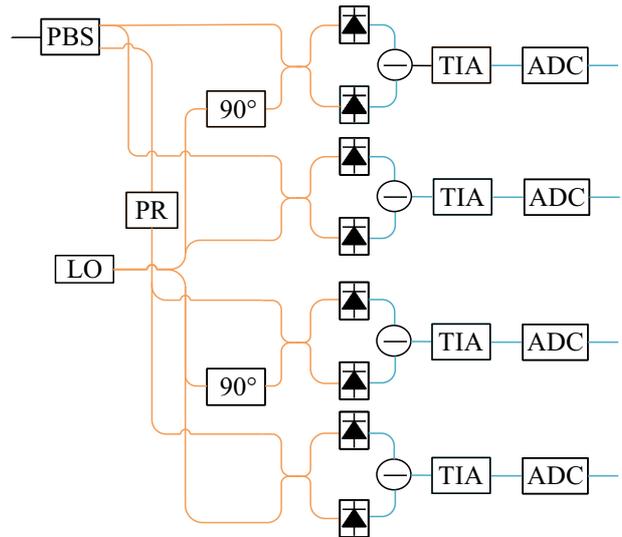
where  $T$  — symbol period. Therefore, the objective of DSP for the received sequence  $y_k$  is to compensate the  $h$  distortions so that the signal could be represented as

$$a'_k = a_k + v'(kT_s), \quad (2)$$

where  $v'$  — additive white Gaussian noise of the same intensity as the ASE noise.

This paper is dedicated to the linear DSP with coherent detection of a signal with DP-QPSK modulation format. The existing algorithms of digital processing were reviewed. The DSP circuit was presented, making it possible to compensate most linear distortions arising in the process of signal transmission and reception. The circuit operability was tested experimentally on the signal with symbol speed of 31.6 GBd, which travelled the line with length of up to 4000 km. Therefore, demodulation of 100 G signal is provided.

This paper consists of 3 sections. Section 1 is dedicated to description of DSP structure and review of solutions which may be used to compensate a certain distortion. Section 2 provides the original implementation of the full tract of digital processing with specification of the used algorithms and their parameters. Validation of operation of the proposed DSP circuit implemented on the actual line with length of up to 4000 km is presented in Section 3. The main results and conclusions of the work done are presented in Section 3.



**Figure 1.** Coherent receiver diagram. PBS — polarization beam splitter, LO — local oscillator,  $90^\circ$  — phase shift by  $90^\circ$ , PR — polarization rotator by  $90^\circ$ , TIA — transimpedance amplifier, ADC — analog-to-digital converter.

In the paper we will use the following agreements in the equations:

- bold font and straight letters — column-vector, uppercase letters — scalar values;
- for vectors the lowercase letters — time area, capital letters — frequency area. The same letters mean the same value in different domains;
- $F[\ ]$  — discrete Fourier transform (DFT),  $F^{-1}[\ ]$  — inverse discrete Fourier transform (IDFT);
- $\times$  — matrix product;
- $\circledast$  — cyclic convolution;
- $\cdot$  — element by element multiplication for vectors.

## 1. DSP structure

An optical signal in a coherent receiver is divided into two components by a polarization beam splitter, and each of them is divided into two quadrature components using  $90^\circ$  hybrid. Four components of the optical signal are then detected by differential photodetectors, and four electric signals are sampled by ADC. Discrete samples from the outputs of four ADCs arrive to the input of DSP, where they are processed by blocks. The diagram of such receiver is shown in Fig. 1.

### 1.1. Automatic gain control

The first block of DSP carries out automatic gain control (AGC) to stabilize the average capacity of the digitized signals.

There are two main types of AGC: the first processes the signal regardless of previously received samples [13], the second takes into account the prehistory of the signal [14].

Let us consider one of the algorithms of the second type — capacity assessment with partial summation (CAPS). The algorithm is classic from the point of view of its reliance on the signal capacity, however its implementation having high degree of parallelism is original and according to our data has not been published previously. The algorithm breaks the arriving signal into blocks of serial samples, assessing the gain ratio for each block. A certain, always the same, quantity of samples and the previous block gain ratio arrive to the algorithm input. For each sample that arrived to the block input, the instantaneous capacity is assessed. Then the algorithm assesses the gain ratio for this block, summing, by analogy with the geometric progression, the previous gain ratio and assessed instantaneous capacities multiplied by members of a certain geometric progression.

The signal samples are processed by the blocks, the size of which determines the algorithm parallelism. A sequence of 4-dimensional real-valued vectors  $\mathbf{x}_{n,j}$  are supplied to the input, which represent samples from four channels of ADC, where  $n$  — block number,  $j$  — serial number of the sample in the block. Then the  $N_{n,j}$  signal capacity at this moment of time may be assessed using equation

$$N_{n,j} = \frac{\mathbf{x}_{n,j}^T \times \mathbf{x}_{n,j}}{2}. \quad (3)$$

The average  $P_{n+1}$  capacity for  $n + 1$ th block is assessed as

$$P_{n+1} = \sum_{j=1}^K \alpha^{K-j} N_{n,j} + P_n \alpha^K, \quad (4)$$

where  $K$  — number of samples in each block, and  $\alpha$  — a certain number from  $(0, 1)$  interval, interpretation of which will be provided below. Sequentially expressing  $P_k$  via  $P_{k-1}$  in equation (4) for  $k = n - 1, \dots, 1$ , we obtain

$$P_{n+1} = \sum_{k=1}^{nK} \alpha^{nK-k} N_k, \quad (5)$$

where  $N_k$  — capacity of  $k$ -th sample in the entire signal. Equation (5) is an analog of the geometric progression. From this equation it follows that the closer  $\alpha$  is to one, for the higher quantity of the previous blocks the average capacity is taken into account, and vice versa, the closer  $\alpha$  is to zero, the fewer blocks contributed to the assessment of the average signal capacity for this block.

The signal is normalized using the following equation

$$\mathbf{y}_{n,j} = G(\alpha, n) \mathbf{x}_{n,j}, \quad (6)$$

where the gain ratio  $G(\alpha, n)$  for each block is defined as follows

$$G(n, \alpha) = \sqrt{\frac{S(\alpha)}{P_n}}, \quad (7)$$

where  $S(\alpha) = \frac{P_0}{1-\alpha}$  — sum of the infinite decreasing geometric progression with denominator  $\alpha$ , and initial member  $P_0 = 1$ .

In the ideal case, when the capacity of all samples is the same and is equal to one, the average capacity of each block, with a rather high serial number, would be equal to  $S(\alpha)$ . Therefore, if  $P_n > S(\alpha)$ , i.e. higher than the needed value, then  $G(n, \alpha) < 1$  — decrease of the input signal value, and vice versa, if  $P_n < S(\alpha)$ , then  $G(n, \alpha) > 1$ .

The above algorithm has two advantages: the samples in one block may be processed in parallel, and it is easy to implement the integer-value analog of the CAPS algorithm.

Let us consider another AGC algorithm described in [15].

The gain ratio at this moment of time  $g(t)$  is specified by equation

$$g(t) = \frac{y(t)}{x(t)} = \exp \left[ K \int_0^t (Y_{ref}^2 - y^2(t)) dt \right], \quad (8)$$

where  $Y_{ref}$  — a constant compliant with the target level,  $x(t)$  — a signal,  $y(t)$  — a normalized signal,  $K = \frac{1}{2\tau Y_{ref}^2}$ ,  $\tau$  — a time interval, when the current capacity is assessed.

Then using (8) we obtain the discrete analog of this ratio

$$g_k = \exp(z_k), \quad (9)$$

where  $z_k$  — integral value. Let us assess  $z_k$  using trapezoid rule with  $T_s$  pitch value equal to sampling period

$$z_k = KT_s \sum_{n=0}^{k-1} (Y_{ref}^2 - y_n^2). \quad (10)$$

Due to additivity of a certain integral,  $z_k$  may be calculated recurrently

$$z_k = z_{k-1} + KT_s [(Y_{ref}^2 - y^2(k-1))]. \quad (11)$$

Since the exponent factors are summed up with exponent multiplication,  $g_k$  may also be calculated recurrently

$$g_k = g_{k-1} \exp [KT_s (Y_{ref}^2 - y_{k-1}^2)]. \quad (12)$$

The above algorithm has two substantial disadvantages: calculation of each element calculates the exponent, which requires more resources, and the algorithm also normalizes each sample sequentially, which substantially complicates its paralleling.

Therefore, AGC is a critical block, without which the operation of the following DSP blocks is not possible. AGC algorithms processing the signal by blocks operate faster compared to recursive algorithms, but at the same time, as well as the recursive ones, take into account the previously processed signal, which makes such algorithms adaptive.

## 1.2. Chromatic dispersion (CD)

CD is a dependence of the effective group refraction index on the optical carrier frequency, which causes distortion of the signal temporal shape [16]. In particular, optical impulses corresponding to the transmitted symbols, expand and overlap with the neighboring symbols. Since a telecommunication signal has bandwidth of several dozens GHz and spreads along the fiber to hundreds of kilometers, the impact of such distortion turns out to be substantial.

### 1.2.1. CD assessment methods

To compensate dispersion, it is necessary to first assess its value. There are two essentially different approaches, namely assessment using additional information and blind assessment.

Most non-blind assessments are based on fractional Fourier transform (FFT) [17], they differ only by the known sequence used to assess CD. It may be a sequence with linear frequency modulation [18], a sequence of certain QPSK transitions of the signal [19] or a sequence with constant amplitude zero autocorrelation CAZAC [20]. Regardless of the proposed sequence, all of these papers use FFT, then the integral is calculated from the module of the fourth degree of the signal after FFT, and, finally, depending on the used sequence, the optimal value of the FFT order is found. This optimal value is used to calculate the CD value. Let us note several disadvantages of such approach, namely: the need to measure the reference order of FFT at the known dispersion; high computational complexity. For example, to implement the method described in [18], a lengthy sequence is required of 2048 known symbols. It is worth noting the paper [21], the algorithm in which operates using the principle similar to the method presented in [18], however, it does not add a special signal, but uses the fact that the signal with CD is chirped in the frequency area.

Most blind assessments are based on the preliminary compensation of CD, the value of which is sorted until a certain target function is optimized [22–26]. A target function used is signal capacity [22,23], or a timing synchronization error computed by different methods [24–26]. A disadvantage of blind methods is obviously the fact that they are based on sorting. For example, we must compensate dispersion 45000 ps/nm with accuracy of up to 300 ps/nm, then in order to sort all values with such pitch, ~150 sorting iterations are required. The simplest metric of compensation correctness is the convergence of the entire system, therefore, in each of the iterations it is necessary to compensate the chromatic dispersion and to obtain the converging system at the correct value.

The blind assessment may also be performed by searching for the second autocorrelation peak [27], however, the disadvantage of such method would be high computational complexity, while for the reliable peak detection at high dispersion it is necessary to process a sequence of long length, for example, 32768 samples in [27].

A separate class of assessors may include those that are based on the measurement of fiber transmittance matrix  $H_f(f)$  in the spectral representation [28]. Let us designate the inverse matrix via  $W(f)$ :

$$W(f) = H_f^{-1}(f). \quad (13)$$

Then the argument of the dispersion operator of  $H_{CD} = \exp(jf^2\beta_2 2\pi^2)$  type may be assessed using the following equation:

$$\arg(H_{CD}^{-1}) = \arg(\sqrt{\det(W(f))}) = -f^2\beta_2 2\pi^2. \quad (14)$$

More details about the transmittance matrix assessment methods are given in Section 1.4, however, here we will give an example of the paper, where  $H_f(f)$  matrix is assessed specifically for assessment of CD. The fact that the matrix must be known prior to dispersion assessment results in certain limitations for the DSP circuit. Most methods of timing synchronization and compensation of the carrier frequency deviation either do not work in the presence of CD, or the accuracy of their work is low, which also results in the fact that the assessment of the fiber transmittance matrix turns out to be incomplete. In paper [29] it is suggested to assess the channel using complementary Golay sequences [30], however, the length of such sequences turns out to be too long (256 symbols [29]).

### 1.2.2. CD compensation

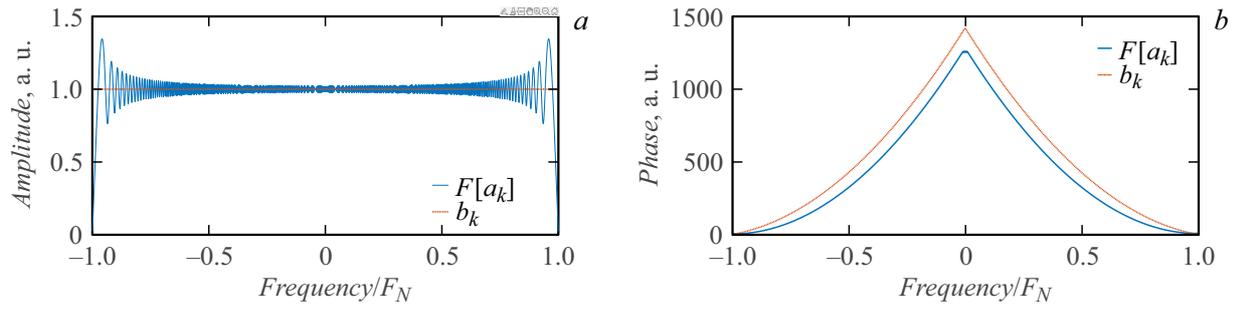
After the CD is assessed, its effect may be compensated. The main idea of CD compensation is using an inverse operator, which was obtained analytically from the NSE (nonlinear Schrodinger equation) [16]. In reality many effects operate jointly, and the dispersion effects may only be roughly compensated in this manner. However, such accuracy is sufficient, since the compensation of the residual dispersion may be undertaken by the subsequent algorithms of digital processing (for example, an adaptive filter).

An obvious question arises immediately: in which area is CD to be compensated, in the time or in the frequency one? First let us consider the compensation in the time area. In this case the filter built on the inverse impulse response of the CD operator appears as [31]:

$$a_k = \sqrt{\frac{jcT_s^2}{D\lambda^2z}} \exp\left(-j\frac{\pi cT_s^2}{D\lambda^2z}k^2\right), \\ -\left[\frac{N_{CD}}{2}\right] \leq k \leq \left[\frac{N_{CD}}{2}\right], \quad (15)$$

where  $c$  — light velocity,  $D$  — CD coefficient,  $\lambda$  — central wavelength of the optical signal,  $z$  — fiber length,  $N_{CD}$  — length of the filter impulse response, which defines the number of signal samples that the CD is compensated for.

However, for formula [15] there is a limitation of its potential length  $N_{CD}$ . The reason for its occurrence is aliasing. It is caused by exceeding the frequencies of the



**Figure 2.** Comparisons of AFR (a) and PFR (b) for  $F[a_k]$  from formula (15) and  $b_k$  from formula (19). Filter parameters:  $D = 16$  ps/(nm·km),  $\lambda = 1550$  nm,  $F_s = 64$  GHz,  $z = 3100$  km,  $N_F = N^A = 1629$ ,  $F_N = F_s/2$  — Nyquist frequency.

Nyquist frequency filter  $\omega_N = \frac{\pi}{T_s}$  [32] ( $-j\phi_k$  — exponent index in formula (15)):

$$|\omega_k| = \left| \frac{d\phi_k}{dt} \right| = \{t = T_s k\} = \left| \frac{2\pi c T_s}{D\lambda^2 z} k \right| < \frac{\pi}{T_s}. \quad (16)$$

It is easy to derive the limitation of the maximum number of filter coefficients before aliasing from this:

$$N^A = 2 \left\lfloor \frac{|D|\lambda^2 z|}{2cT_s^2} \right\rfloor + 1, \quad (17)$$

where  $\lfloor x \rfloor$  means that rounding  $x$  to the nearest integer is lower than  $x$ .

Apart from the upper limit, there is also a lower limit for the potential length  $N_{CD}$ . Impact of CD on the signal results in the considerable expansion of the signal impulses [16]. Thereby to recover at least one impulse, it is necessary to use information from the adjacent impulses, and their quantity depends on the CD value [16] and the symbol succession period  $T$ :

$$N^B = 2 \left\lceil \sqrt{1 + \left( \frac{\pi |D|\lambda^2 z|}{8c(\ln 2T)^2} \right)^2} \right\rceil + 1, \quad (18)$$

where  $\lceil x \rceil$  means that rounding  $x$  to the nearest integer is higher than  $x$ . Value  $N^B$  is obtained in approximation of Gaussian impulses, which does not correspond to reality, where usually the impulses of the raised cosine shape are used. Therefore, value  $N^B$  is only suitable as assessment.

Sampling rate  $F_s$  and symbol speed  $R_s$  are connected via a value that defines the number of points after ADC per one symbol  $OS = F_s/R_s$ . According to the sampling theorem [33], sampling rate must exceed the symbol speed at least twice  $F_s \geq 2R_s$ , in order to avoid aliasing. In practice the number of points per symbol is chosen as slightly above two. If this fact is used, it turns out that  $N^B < N^A$ , which means that the compensation in the time area is possible.

Let us now consider the compensation operator in the frequency area, which appears as [31]:

$$b_k = \exp \left[ -j \frac{D\lambda^2 z}{\pi c} \left( \frac{k}{N_F} \omega_N \right)^2 \right], \quad (19)$$

$$-\frac{N_F}{2} \leq k \leq \frac{N_F}{2} - 1,$$

where  $N_F$  — number of signal samples, where Fourier transform must be taken to multiply by  $b_k$ . Dimension  $N_F$  is only limited at the bottom by value  $N^B$  and is not limited at the top.

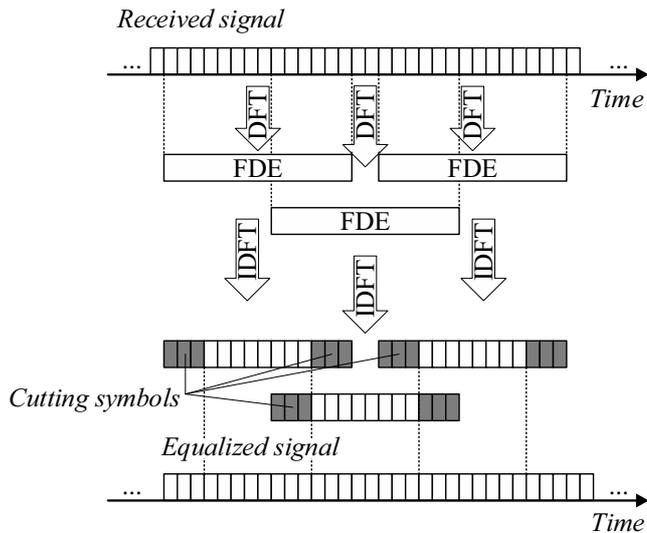
As one may see, there are no issues with aliasing in formula (19), as the maximum frequency of the operator is limited by Nyquist frequency  $\omega_{FD}^{\max} = \omega_N$ . For the time operator the limitation is determined by value  $N^A$ , which provides the maximum frequency  $\omega_{TD}^{\max} = \frac{2\pi c}{D\lambda^2 z} \lfloor N^A \rfloor T_s$  [31].

Fig 2 shows AFR and PFR of filters in the time and frequency areas. As can be seen, the filter built in time has substantial distortions in the frequency area. This is related to the short length of Fourier transform, which is limited by aliasing ( $N^A$ ).

There is one more argument for the benefit of dispersion compensation in the frequency area. Starting from a certain filter length, the signal convolution with the time filter turns out to be a more capacity-demanding operation vs Fourier transform, member by member multiplication by frequency filter and inverse Fourier transform [34].

As discussed, to compensate CD in one sample, information is necessary from at least  $N^B$  samples. Besides, the dispersion will fully be compensated only in the sample that is located in the middle. To compensate dispersion in the next sample, the signal sequence will have to be shifted, and all the operations will have to be repeated. Therefore, to compensate the dispersion, a block is taken from  $N_{CD}$  samples, where only the central part of the samples goes to the output, and the next block is taken as overlapping with the next one. Fig. 3 shows the diagram proposed in paper [35].

From the above one may conclude that CD compensation in the frequency area is more beneficial. However, even for the frequency area there is a lower limitation for the



**Figure 3.** Diagram of block compensation of CD in frequency area with overlapping neighboring blocks [35].

dimension of filters  $N^B$ , which for higher CD values results in a higher growth of computational complexity of the dispersion compensation block.

### 1.3. Timing synchronization

Algorithms of timing synchronization are aimed at solving two tasks related to sampling of the signal that arrived from the line. First, in practice the sampling rate of ADC  $F_s$  does not match the symbol speed of signal  $R_s$ , multiplied by integer number. With time this difference accumulates, and if at first we were lucky, and the samples were located in the center of the impulse, in a few impulses the samples will get both into gaps between impulses and into neighboring impulses even. If we approximately know both  $F_s$  and  $R_s$ , then we will be able to compute this shift and interpolate the samples. Second, frequencies of the transmitter (DAC) and receiver (ADC) are stable with certain accuracy. For this reason we have to assess the ratio between  $F_s$  and  $R_s$  in real time.

The objective of the timing synchronization block consists in setting the required integer number of samples per symbol after its operation, besides, one of these samples must be located exactly in the center of symbol intervals. Delay (advance) of the sample relative to the optimal position is called an error, to assess which there are many known algorithms, which may be divided into two classes: with inverse and direct association. In algorithms with inverse association, to calculate an error, samples are used after compensation of such error, and for algorithms with direct association — before compensation. The main representatives of the first class are Muller–Muller algorithm [36], Gardner algorithm [37,38], of the second one — Oerder–Meyer algorithm [39], Lee [40] and Godard algorithms [41]. Let us note at once that all given algo-

gorithms are poorly sensitive to the carrier phase deviation, however require the preliminary compensation of CD to the level below 300 ps/nm and rough compensation of carrier frequency offset to dozens of MHz. Now let us discuss in more detail some algorithms of timing synchronization.

#### 1.3.1. Godard algorithm

One of the outstanding representatives of the class of algorithms with direct association is Godard algorithm [41]. The task in the original article is set as follows: let the signal after the matched filter represents

$$y(t) = \sum_n a_n h_{TxF}(t - nT) \exp(j2\pi f_0 t) + v(t), \quad (20)$$

where  $\{a_n\}$  — transmitted data,  $h_{TxF}(t)$  — full impulse response of the system to the circuit of coherent reception (transmitter and fiber),  $f_0$  — carrier frequency,  $v(t)$  — white noise filtered by matched filter.

This signal is sampled at frequency  $F_s = R_s = 1/T$ , however, sampling is accompanied by a certain time offset  $\tau_s$ :  $\{y(\tau_s + kT)\}$ . It is worth noting that the presence of offset is equivalent to sampling with frequency different from  $R_s$ . It is necessary to define  $\tau_s$ , maximizing the energy of the received signal. The Godard’s maximum is found by stochastic gradient descent.

Having taken Fourier transform from formula (20) and identified all components in it dependent on  $\tau_s$ , the explicit form of metric was obtained for maximization

$$\epsilon^2(\tau_s) = T \int_{f_0}^{f_0+1/T} |H_s(f, \tau_s)|^2 df, \quad (21)$$

where  $|H_s(f, \tau_s)|$  is determined by equation

$$|H_s(f, \tau_s)| = \frac{1}{T} \sum_k H_{TxF} \left( f - f_0 - \frac{k}{T} \right) \exp \left( j2\pi \left( f - \frac{k}{T} \right) \tau_s \right). \quad (22)$$

As a result, Godard obtained the explicit formula for the gradient descent pitch, i.e. for the value, by which it is necessary to adjust  $\tau_s$ :

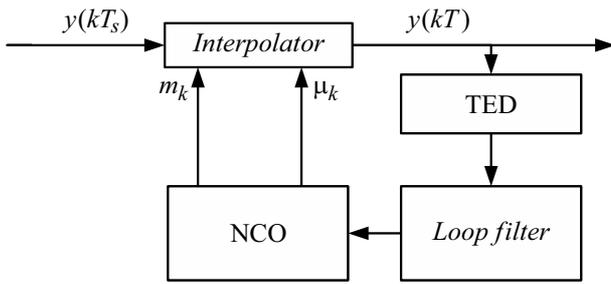
$$p_n = \text{Im} (g_1^*(nT + \tau_s) g_2(nT + \tau_s)), \quad (23)$$

where  $g_1$  — received signal that passed via low pass filter with center in  $f_0 - 1/2T$ , and  $g_2$  — via filter with center in  $f_0 + 1/2T$ .

And the tune-up itself  $\tau_s$ :

$$\tau_s(n+1) = \tau_s(n) - \mu_1 p_n - \mu_2 \sum_{k=1}^n p_k, \quad (24)$$

where  $\mu_1, \mu_2$  — coefficients of a proportionately integral regulator.



**Figure 4.** Circuit diagram of Gardner timing synchronization from [43,45]. NCO (number controlled oscillator) — oscillator that defines the position of points for interpolation, TED — timing error detector.

Usually instead of formula (21) a discrete analog is used (25) [42]:

$$\epsilon_{disc}^2 = \sum_{k=0}^{N/2-1} \text{Im} \left( X_{in}^{ret} X_{in, k+N/2}^{ret*} \right), \quad (25)$$

where  $X_{in}^{ret}$  — Fourier transform of data at the input to the timing synchronization block,  $\epsilon_{disc}$  — discrete analog of the metric for maximization.

### 1.3.2. Gardner algorithm

Among algorithms with inverse association, Gardner algorithm may be identified [37,43,44]. The overall circuit diagram of digital timing synchronization for algorithms with inverse association is shown in Fig. 4.

The error detector is Gardner detector from [37]:

$$e(k) = I(k - 1/2)[I(k) - I(k - 1)] + Q(k - 1/2)[Q(k) - Q(k - 1)], \quad (26)$$

where  $I(k), Q(k)$  — real and imaginary parts of the signal  $y(kT)$  after the interpolator. A feedback filter is a proportionately integral regular, the signal from which goes to the block of NCO (numerically controlled oscillator). NCO is responsible for conversion of error in  $m_k$  — number of sample for interpolation,  $\mu_k$  — where to from  $m_k$  it is necessary to interpolate.

Paper [46] provides spectral representation of the Gardner error detector, and its modification to handle impulses close to Nyquist ones. Implementation of the Gardner algorithm in parallel form is considered in article [47].

According to article [42], timing synchronization error assessment algorithms are not very different in efficiency, therefore the selection of a certain algorithm depends on the convenience of designing a specific system.

### 1.3.3. Interpolator

After the symbol speed or error were assessed in time, it is necessary to obtain 2 points per symbol. For this purpose it is suggested to use interpolation. Farrows

interpolators are used most often [48], which are suitable for implementation in VLSIC (very-large-scale integrated circuits). As an example, let us consider one most often used cubic interpolator:

$$c_1 = -\frac{1}{6}\mu_k^3 + \frac{1}{6}\mu_k,$$

$$c_2 = \frac{1}{2}\mu_k^3 + \frac{1}{2}\mu_k^2 - \mu_k,$$

$$c_3 = -\frac{1}{2}\mu_k^3 - \mu_k^2 + \frac{1}{2}\mu_k + 1,$$

$$c_4 = \frac{1}{6}\mu_k^3 + \frac{1}{2}\mu_k^2 + \frac{1}{3}\mu_k,$$

$$x_{out}^{ret}(kT_i) = \sum_{n=1}^4 c_n \cdot x_{in}^{ret}(m_k + n - 3), \quad (27)$$

where  $x_{in}^{ret}$  — input of timing synchronization block,  $x_{out}^{ret}$  — output of timing synchronization block.

## 1.4. Adaptive filtration

Many distortions change in time, such as PMD or polarization rotation. To compensate distortions of such type, an adaptive filter (AF) is used, coefficients of which are tuned in time. In process of signal propagation along the fiber, signal polarizations are mixed because of random birefringence in the fiber. This transform may be described using Jones matrices [49], therefore, in order to compensate such mixing, the classic AF structure is also built in the form of matrix with size  $2 \times 2$ .

### 1.4.1. Filter structure

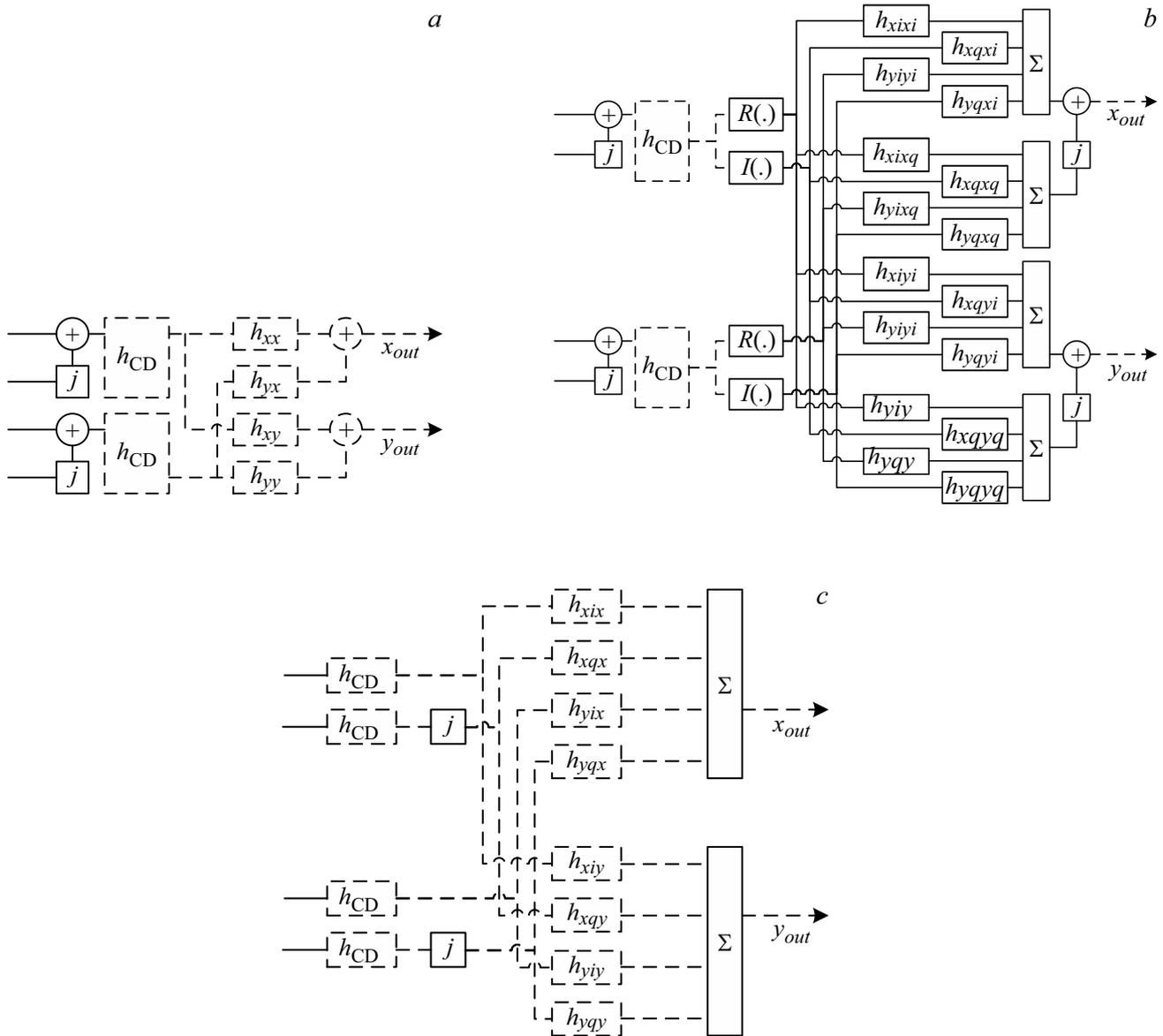
Let us consider AF in the time area  $2 \times 2$  with two complex inputs and two complex outputs [32]. Its structure is shown in Fig. 5, a.

Mathematically the impact of filter  $2 \times 2$  is expressed in the following equations (28):

$$\begin{aligned} x_{out}[k] &= \mathbf{h}_{xx}^T \times \mathbf{x}_{in}[k] + \mathbf{h}_{yx}^T \times \mathbf{y}_{in}[k], \\ y_{out}[k] &= \mathbf{h}_{xy}^T \times \mathbf{x}_{in}[k] + \mathbf{h}_{yy}^T \times \mathbf{y}_{in}[k], \end{aligned} \quad (28)$$

where  $\mathbf{h}_{xx}, \mathbf{h}_{yx}, \mathbf{h}_{xy}, \mathbf{h}_{yy}$  — column-vectors of AF coefficients,  $\mathbf{x}_{in}[k], \mathbf{y}_{in}[k]$  —  $k$ - complex input vectors into filter, each made of two quadratures,  $x_{out}[k], y_{out}[k]$  —  $k$ - filter outputs. Formulas for other MIMO structures are obtained by analogy.

The structure of the classic AF  $2 \times 2$  may be changed, having taken into account the fact that each of the signal polarizations represents two independent quadratures. In article [50] it is stated that AF structure  $4 \times 2$ , shown in Fig. 5, c in contrast to the other ones given in Fig. 5, is resistant to delays between quadratures arising on the receiving part. The reason for its resistance consists in the fact that the dispersion operator in this case does not mix



**Figure 5.** MIMO (multiple input multiple output) circuits from [50] complex  $2 \times 2$  (a), real  $4 \times 4$  (b), complex  $4 \times 2$  (c). Solid — real numbers, dash — complex numbers,  $h_{CD}$  — operator of accumulated chromatic dispersion compensation.

the signals arriving to the input of AF. For example, the classic structure  $2 \times 2$  for DP-16QAM modulation format gives a penalty of around 2dB by SNR (signal-to-noise ratio) for delays of 13% of the symbol, and  $4 \times 2$  shows operation without a penalty [50]. Besides, the circuit of  $4 \times 4$  type (Fig. 5, b) also has a penalty caused by delays between quadratures, since prior to operation of AF the signal quadratures are mixed to compensate CD. Delays between quadratures may vary from a device to a device and in time, therefore it is important to compensate them automatically.

**1.4.2. Target function**

AF coefficients are calculated based on minimization of

a certain target function. There are many types of target functions, one may identify those among them that take into account the signal module only, and those that take into account both module and phase.

There are two different approaches to calculation of the target function values: 1) blind — without reference to known data, 2) non-blind — according to the previously known service sequence.

The most effective blind algorithm is the constant modulus algorithm — CMA [51]. The error for polarization  $p$  ( $p$  — takes on values from the set  $\{x, y\}$ ) may be expressed using equation:

$$\epsilon_p^{CMA} = (1 - |E_p^{out}|^2), \tag{29}$$

where  $E_p^{out}$  — output of the adaptive filter for polarization  $p$ . The square of error CMA  $(\varepsilon_p^{CMA})^2$  is the target function.

If the known subsequence is present in the data, most often the least-square method — LS is used, or the minimum mean square error — MMSE, such solutions in the frequency area are given, for example, in [52]. Let us represent the received signal  $R_x$  and  $R_y$  at frequency  $f_i$  by the equation from the formula (30).

$$\begin{bmatrix} R_x(f_i) \\ R_y(f_i) \end{bmatrix} = \begin{bmatrix} H_{xx}(f_i) & H_{xy}(f_i) \\ H_{yx}(f_i) & H_{yy}(f_i) \end{bmatrix} \begin{bmatrix} T_x(f_i) \\ T_y(f_i) \end{bmatrix} + \begin{bmatrix} No_x(f_i) \\ No_y(f_i) \end{bmatrix}, \quad (30)$$

where  $H_{xx}, H_{xy}, H_{yx}, H_{yy}$  — corresponding components of the channel,  $T_x, T_y$  — spectra of sent signals,  $No_x, No_y$  — noise spectra.

It is possible to divide  $T_x$  and  $T_y$  into two components:

$$\begin{aligned} T_x[f_i] &= S_x[f_i] + M_x[f_i], \\ T_y[f_i] &= S_y[f_i] + M_y[f_i], \end{aligned} \quad (31)$$

where  $S_x[f_i]$  and  $S_y[f_i]$  — the DFT of the signal taken in the window that contains only pilot subsequence, and symbols in the load are equal to zero;  $M_x[f_i]$  and  $M_y[f_i]$  — the DFT of the signal taken in the window that contains only symbols in the load, and the pilot subsequence is equal to zero. Then the following formula is valid:

$$\begin{bmatrix} R_x[f_1] & R_y[f_1] \\ R_x[f_2] & R_y[f_2] \\ \vdots & \vdots \\ R_x[f_{N_c}] & R_y[f_{N_c}] \end{bmatrix} = \begin{bmatrix} S_x[f_1] & S_y[f_1] \\ S_x[f_2] & S_y[f_2] \\ \vdots & \vdots \\ S_x[f_{N_c}] & S_y[f_{N_c}] \end{bmatrix} \mathbf{H}^T + \begin{bmatrix} M_x[f_1] & M_y[f_1] \\ M_x[f_2] & M_y[f_2] \\ \vdots & \vdots \\ M_x[f_{N_c}] & M_y[f_{N_c}] \end{bmatrix} \mathbf{H}^T + \begin{bmatrix} No_x[f_1] & No_y[f_1] \\ No_x[f_2] & No_y[f_2] \\ \vdots & \vdots \\ No_x[f_{N_c}] & No_y[f_{N_c}] \end{bmatrix}, \quad (32)$$

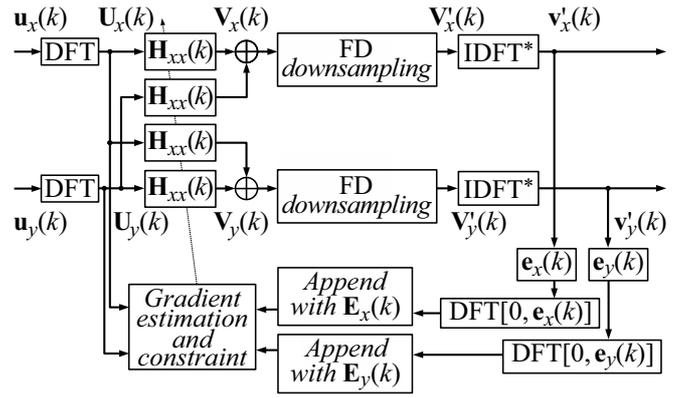
or in a more compact form

$$\mathbf{R} = \mathbf{S}\mathbf{H}^T + \mathbf{M}\mathbf{H}^T + \mathbf{N}_o. \quad (33)$$

Then the LS solution is expressed by equation

$$\hat{\mathbf{H}}_{LS}^T = \arg \min_{\mathbf{H}^T} \text{tr} \left\{ \left( \mathbf{R} - \mathbf{S}\hat{\mathbf{H}}^T \right) \left( \mathbf{R} - \mathbf{S}\hat{\mathbf{H}}^T \right)^* \right\} = \mathbf{S}^\dagger \mathbf{R}, \quad (34)$$

where  $\hat{\mathbf{H}}$  — channel assessment,  $*$  — Hermitian conjugation,  $\mathbf{S}^\dagger = (\mathbf{S}^* \mathbf{S})^{-1} \mathbf{S}^*$  — pseudo-inverse matrix  $\mathbf{S}$ . Besides, article [52] provides the MMSE solution as well. However, even though the MMSE method does provide a more optimal assessment, its implementation requires substantially more computing resources.



**Figure 6.** Circuit diagram of AF from [54]. IDFT\* means that from the entire vector after the inverse discrete transform only the last  $N_{FD}/2$  samples are taken.

### 1.4.3. Minimization method

AF coefficients are updated using a stochastic gradient descent [32]. Let us consider the example of the filter with structure  $2 \times 2$  in the time and frequency areas.

In the time area the coefficients of the filter with CMA metric are updated according to the following rule [32]:

$$\begin{aligned} \mathbf{h}_{xx} &= \mathbf{h}_{xx} + \mu \cdot \varepsilon_x \cdot x_{out}[k] \cdot \mathbf{x}_{in}^*[k], \\ \mathbf{h}_{xy} &= \mathbf{h}_{xy} + \mu \cdot \varepsilon_y \cdot y_{out}[k] \cdot \mathbf{x}_{in}^*[k], \\ \mathbf{h}_{yx} &= \mathbf{h}_{yx} + \mu \cdot \varepsilon_x \cdot x_{out}[k] \cdot \mathbf{y}_{in}^*[k], \\ \mathbf{h}_{yy} &= \mathbf{h}_{yy} + \mu \cdot \varepsilon_y \cdot y_{out}[k] \cdot \mathbf{y}_{in}^*[k], \end{aligned} \quad (35)$$

where  $\varepsilon_x, \varepsilon_y$  — errors of signal CMA metrics,  $\mu$  — convergence speed.

In the frequency area the coefficients of the filter with CMA metric are updated in several stages, since the target function is calculated in the time area, and coefficients are updated in the frequency one. The AF concept in the frequency representation is specified in article [53], where the authors suggested the method of working with the signal having two samples per symbol. Later on in the conference theses [54] the same authors suggested a modification, which made it possible not to divide into even and odd samples and to work with the rational number of samples per symbol. The more recent implementation of AF makes it possible to cut in half the required number of adaptive filters compared to the version from [53]. Because of the above advantages, we will consider only the more recent version.

AF circuit diagram is shown in Fig. 6. Time sequence  $u_{x,y}(n)$  is supplied to the input for X and Y polarizations accordingly. Samples are taken more frequently than symbols, rational number of times,  $n$  — sample index. Further without loss of generality, an example with two points per symbol will be considered. For effective calculation of convolutions, the overlap-saved method will be used with overlap factor 50% (article [53] states that this is the

most optimal value). Therefore, the considered vector in the frequency space will look as follows:

$$\begin{aligned} \mathbf{U}_{x,y}(k) &= F [u_{x,y}(kN_{FD} - N_{FD}), \dots, u_{x,y}(kN_{FD} + N_{FD} - 1)]^T, \end{aligned} \tag{36}$$

where  $k$  — block index,  $\mathbf{U}_{x,y}(k)$  — DFT of  $k$ -th input vector  $\mathbf{u}_{x,y}(k)$ ,  $N_{FD}$  — half of length of the input vector.

The AF equation appears as follows:

$$\begin{aligned} \mathbf{V}_x(k) &= \mathbf{H}_{xx}(k) \cdot \mathbf{U}_x(k) + \mathbf{H}_{xy}(k) \cdot \mathbf{U}_y(k), \\ \mathbf{V}_y(k) &= \mathbf{H}_{yx}(k) \cdot \mathbf{U}_x(k) + \mathbf{H}_{yy}(k) \cdot \mathbf{U}_y(k), \end{aligned} \tag{37}$$

where  $\mathbf{H}_{xx}(k)$ ,  $\mathbf{H}_{xy}(k)$ ,  $\mathbf{H}_{yx}(k)$ ,  $\mathbf{H}_{yy}(k)$  —  $k$ - values of the corresponding AF vectors. Further sampling is decreased in the frequency area:

$$\begin{aligned} \mathbf{V}'_{x,y}(k) &= (\mathbf{V}_{x,y}(k) + \mathbf{V}_{x,y}(k + N_{FD}))/2, \\ \mathbf{v}'_{x,y}(k) &= \text{latest } N_{FD}/2 \text{ report} / F^{-1} [\mathbf{V}'_{x,y}(k)], \end{aligned} \tag{38}$$

$\mathbf{v}'_{x,y}(k)$  — output vector in time with one sample per symbol. It should be noted that article [54] appears to contain an error, its averaging of  $\mathbf{V}'_{x,y}(k)$  does not include division by 2, and it is necessary.

Then the error vector  $\mathbf{e}_x(k)$  is calculated for polarizations  $x$  and  $y$  with size  $N_{FD}/2$ :

$$\mathbf{e}_{x,y}(k) = [\mathbf{I}_{N_{FD}/2} - \mathbf{v}_{x,y}(k) \cdot \overline{\mathbf{v}_{x,y}(k)}] \cdot \mathbf{v}_{x,y}(k), \tag{39}$$

where  $\mathbf{I}_{N_{FD}/2}$  — vector of size units  $N_{FD}/2$ .

The error in the frequency area  $\mathbf{E}_{x,y}(k)$  is calculated using the equation

$$\mathbf{E}_{x,y}(k) = \text{DFT} [\mathbf{O}_{N_{FD}/2}; \mathbf{e}_{x,y}(k)], \tag{40}$$

where  $\mathbf{O}_{N_{FD}/2}$  — vector of size zeroes  $N_{FD}/2$ .

Then, in order to use the error on AF length, the vector size increases by repeating

$$\mathbf{E}'_{x,y}(k) = [\mathbf{E}_{x,y}(k); \mathbf{E}_{x,y}(k)]. \tag{41}$$

Then the average gradient is calculated

$$\nabla_{pq}(k) = \text{first } N_{FD} \text{ counts } F^{-1} [\mathbf{E}'_p \cdot \text{conj} \{ \mathbf{U}_q(k) \}], \tag{42}$$

where  $pq \in \{xx, xy, yx, yy\}$ .

And, finally, the AF coefficients are updated

$$\mathbf{H}_{pq}(k + 1) = \mathbf{H}_{pq}(k) + \mu_{FD} \text{DFT} [\nabla_{pq}(k); \mathbf{O}_{N_{FD}}], \tag{43}$$

where  $\mu_{FD}$  — coefficient of convergence of the stochastic gradient descent.

AF coefficients must be initialized for any target function, however, for CMA, which has a singularity problem [55], search for optimal initial coefficients becomes a separate task [56].

### 1.5. Carrier frequency deviation

Coherent reception requires mixing of a signal from the line with a reference laser at the frequency equal to the carrier frequency of the signal. This method is called a homodyne approach and requires significant efforts to synchronize frequencies and phases of lasers on the transmitter and the receiver. In practice intradyne reception is used, when frequency offset between lasers is not exactly equal to zero, but may vary within several GHz. The payment for the intradyne reception of the signal is addition of the deviation of laser frequency on the receiver from the laser frequency on the transmitter to the list of distortions. Such deviation results in the fact that phase distortion is added to the signal that quickly changes in time.

#### 1.5.1. Methods to assess deviation of carrier frequency

If only distortions caused by lasers remain by the moment of frequency assessment, the signal at the input to the block appears as

$$y'[k] = a_k \exp(i(\varphi[k] + 2\pi\Delta f k T)), \tag{44}$$

$\varphi[k]$  — phase distortions of laser,  $\Delta f$  — frequency offset.

Assessment may be carried out in the time area, using the fourth degree of generating two subsequent reports [57]. It can be shown that for QPSK:

$$(y'[k]y'[k-1])^4 \propto \exp(4i\Delta\varphi[k]). \tag{45}$$

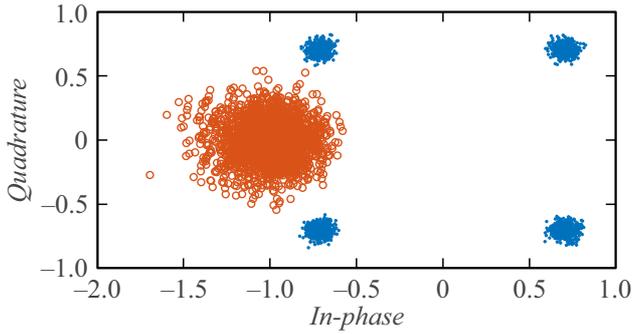
In absence of additive noise  $4\Delta\varphi$  has round Gaussian distribution provided for by the phase noise of the laser with average value  $8\pi\Delta f T$ , so that the probability-density function (PDF) appears as [57]:

$$f(4\Delta) = \frac{\exp(\kappa \cos(4\Delta\varphi - 8\pi\Delta f T))}{2\pi I_0(\kappa)}, \tag{46}$$

where  $\kappa$  is related to the width of the laser line. Using the received PDF, we may assess the parameter of interest by maximum likelihood method [58], which assesses  $\Delta f$  as follows:

$$\Delta f = \frac{1}{8\pi T} \arg \left\{ \sum_{k=1}^N (y'[k]y'^*[k-1])^4 \right\}. \tag{47}$$

The main advantage of the method is its implementability in the time area. However, it has high computing complexity because of the fourth degree of product. This is also a large problem in integer number implementation. The main disadvantage of the method are the conditions of its applicability, namely the fact that it should be found after timing synchronization and adaptive filtration. In practice the assessment of shift may be located before these blocks as well, however, it requires using a vary large averaging, much larger than for the method which will now be considered next.



**Figure 7.** Assessment of carrier frequency deviation by fourth degree of signal. Blue is a QPSK-signal, and orange is its fourth degree.

The problems of large averaging regarding the computational complexity may be mitigated by changing the order of operations and removing averaging beyond the argument function [57]:

$$\Delta f[k] = (1 - \mu)\Delta f[k - 1] + \mu \frac{\arg \left\{ (y'[k]y'^*[k - 1])^4 \right\}}{8\pi T}. \quad (48)$$

Another popular method is assessment of frequency deviation in the frequency area. It is implemented by searching for a peak in the spectrum of the fourth degree of signal [59]:

$$\Delta f[n] = \frac{1}{4} \arg \left( \max_{f \in [-1/2, 1/2]} \sum_{p \in x, y} \mathbf{f}_p(f) \right), \quad (49)$$

$$\mathbf{f}_p(f) = F \left[ \mathbf{y}_p^4[n] \right],$$

where  $\mathbf{y}_p[n]$  —  $n$ -th vector of input signal.

The feature of QPSK format is the fact that in the fourth degree of the signal the phase component responsible for information disappears. In Fig. 7 you can see that all four QPSK states (blue points) are pulled together as one (orange circles), which with time will move along the circumference because of frequency offset.

As mentioned already, this method may work before timing synchronization and adaptive filtration, besides, it does not require increase in the length of the processed block. Since the method still uses the fourth degree, it is also costly from the resources point of view, plus it requires a transition from the time to the frequency domain.

As most distortions, frequency offset may be assessed using the known sequence [60,61]. However, such method requires both additional redundancy and special position of known symbols. Besides, change of the laser in time is quite a fast process, around 500 GHz/s for state-of-the-art  $\mu$ ITLA [62] (micro integrated tunable laser assembly), which also imposes limitations upon the structure of the pilot sequences.

### 1.5.2. Compensation of deviation of carrier frequency

After the frequency deviation was assessed, it must be compensated. Having calculated the deviation of frequency  $\Delta f[n]$ , it is possible to calculate the phase for the entire block as  $\phi_{\Delta f}[n] = \Delta f[n]\mathbf{k}[n]$ , where  $\mathbf{k}[n]$  — time vector of  $n$ -th block. Besides, it is necessary to remember the last value of the phase from the previous block  $\phi_{\Delta f}^{last}[n - 1]$ . Then, using this phase, the complex exponent is calculated, which is multiplied by signal:

$$\mathbf{y}_p^{rest}[n] = \mathbf{y}_p'[n] \cdot \exp \left( -i2\pi(\phi_{\Delta f}[n] + \phi_{\Delta f}^{last}[n - 1]) \right). \quad (50)$$

The disadvantage of such approach is the need to calculate the complex exponent for each block, and also the fact that the signal should remain in the time domain. However, it makes it possible to accurately compensate the frequency offset.

A more simple from the implementation point of view is the compensation method based on cyclic shift of signal spectrum [52]. The problem of this method consists in the fact that the shift may be performed only for the integer number of samples, therefore the assessment has to be divided into integer and fractional parts:  $\Delta f^{int} = \text{round}(\Delta f)$ ,  $\Delta f^{frac} = \Delta f - \Delta f^{int}$ . For compensation of the fractional part it is necessary to improve the sampling by interpolation, so that a cyclic shift by  $\Delta f^{frac}$  samples may be performed. This approach is very costly in terms of computing complexity. Therefore, the shift may be used only when the accuracy of the entire part of the assessment is sufficient, and the remaining part is compensated by one of the following blocks.

### 1.6. Frame synchronization

Data is sent to the line in a structured manner, it is necessary to find such structure upon reception. Most often the data is broken into frames of the same length, and a service sequence is installed in the beginning of each frame.

The main method to detect such sequences is correlation analysis, which searches for a correlation between the received data and previously known sent sequence. The correlation maximum indicates frame start. Therefore, the sequence must have high autocorrelation properties, besides, it should not interfere with the sent signal noise-like features. Besides, sequences for different polarizations must be orthogonal to each other. The CAZAC sequences have the above properties [63]. The example of system operation with such training sequences is considered in [64]. There are other types of sequences suitable for detection of the start of frame, for example, Barker sequences [65] or Golay sequences [30].

Besides, the literature considers a method to detect the start of frame using Schmidle metric [66,67]. However, it is inferior to the correlation one since the sought-for peak has width, and in presence of noise it may cause detection errors. Advantages of this method may include

the fact that the ratio of the power of detected peak to noise level is higher.

If the sequence length is increased to the values of the order of 1000, one may achieve good correlation peaks even at SNR = 3 dB [68]. In practice the length of the correlation sequence amounts to, for example, 22 symbols [12].

### 1.7. Recovery of carrier phase

Coherent communication uses independent lasers on the transmitter and the receiver. The end width of the laser emission line corresponds to certain noise of the carrier phase. Currently the industry widely uses lasers with width of the line of the order of 100 kHz. To compensate distortions caused by deviation of the carrier phase and all the similar ones, the following types of algorithms are also used:

- blind [69–71];
- non-blind, relying on pilot symbols:
  - as a separate block [72–74];
  - within an adaptive filter [52].

One of the first algorithms for recovery of the carrier phase may be Viterbi algorithm [69]. It is based on the fact that raising to a certain degree makes the information component of the signal disappear, and only the part responsible for the distortion remains. Therefore, observing the behavior of the corresponding degree of the signal, one may track the distortion and compensate it. Let us consider this algorithm in more detail using the example of QPSK modulation. In the case, when only phase distortions of the carrier remained, let us present signal  $y(k)$  as

$$\hat{y}[k] = a_k \exp(i\varphi[k]), \tag{51}$$

$\varphi(k)$  — phase distortion of the carrier. Let us raise the signal to the fourth degree

$$(\hat{y}[k])^4 = a_k^4 \exp(4i\varphi[k]) = \exp(i\pi) * \exp(4i\varphi[k]). \tag{52}$$

Therefore, having averaged the noise in the window with length of  $2N + 1$ , one may find the sought-for phase using the following equation:

$$\varphi[k] = \frac{1}{4} \arg \left\{ \frac{1}{2N + 1} \sum_{n=-N}^N y[k + n]^4 \right\}. \tag{53}$$

Let us also consider the blind algorithm without feedback [71], which does not require raising to the fourth degree. Precise raising to a degree is very demanding to the capacity of the final number, therefore this algorithm turns out to be less demanding for integer-value implementation.

Let us define the auxiliary function receiving at the input  $k$ -th distorted symbol  $y[k]$  and number  $\varphi(b) = \frac{b}{B} \pi$ ,

$b \in \{\lfloor \frac{-B}{2} \rfloor, \dots, \lfloor \frac{B}{2} \rfloor\}$ , where  $B$  — preset constant that defines the accuracy of search for the noise value

$$f(y[k], \phi(b)) := |y[k]e^{i\varphi(b)} - [y[k]e^{i\varphi(b)}]_D|^2, \tag{54}$$

$[A]_D$  means the nearest symbol on the complex plane in relation to  $A \in \mathbb{C}$  point.

The algorithm receives at the input  $k$ -th noisy symbol,  $N$  previous distorted symbols and  $N$  subsequent distorted symbols, and then minimizes the function

$$s_k(b) = \sum_{n=-N}^N f(y[k - n], \phi(b)). \tag{55}$$

Let function (55) reach minimum at  $b = b_0$ , then the result of the algorithm operation will be  $a'_k$  sample:

$$a'_k = \hat{y}_k e^{-i\varphi(b_0)}. \tag{56}$$

The disadvantages of all blind algorithms are the uncertainty of phase of type  $n\pi/2$  (where  $n$  — integer value) and demand for computing resources. It is possible to solve these issues by compensation of phase distortions using phase analysis according to the previously known sequence [72–75]. In the literature it is called a pilot subsequence, the symbols of which follow once per several dozens of symbols (for example, in the recommendation [12] — the pilot repetition cycle is equal to 32 symbols). One of the simplest and yet effective algorithms of finding phase distortion  $\varphi(k)$  based on the pilot subsequence is comparison of phases of the subsequent received pilots with their sent version:

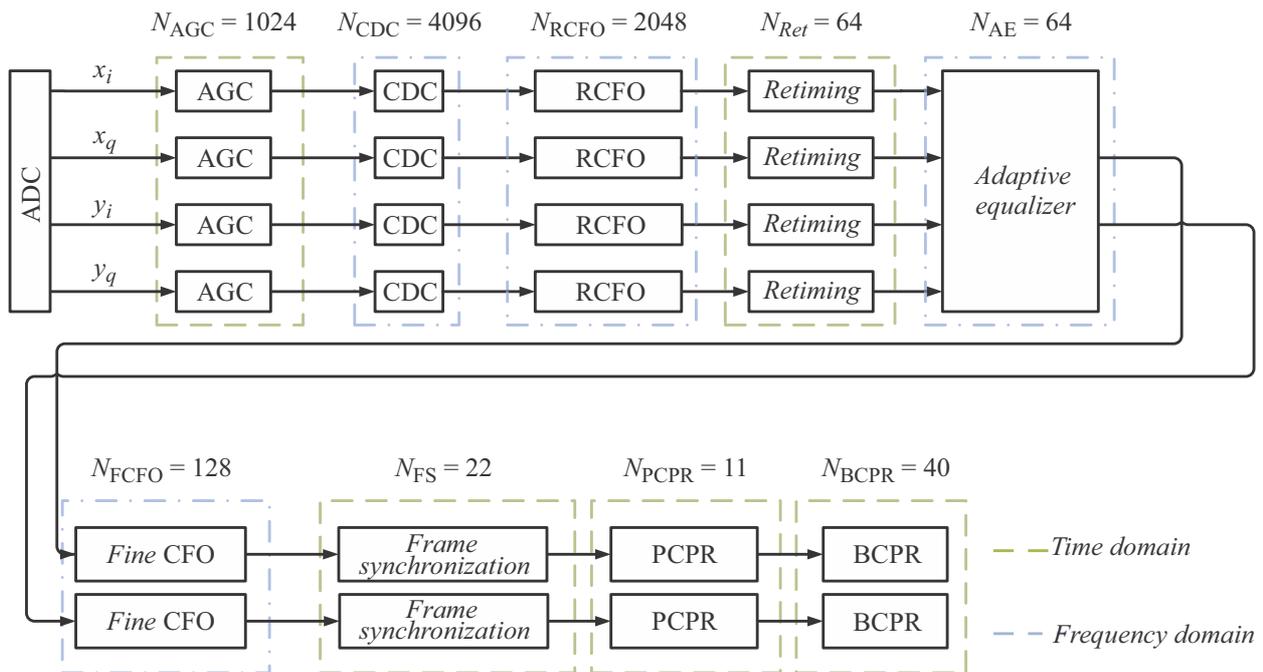
$$\varphi[k] = \arg\{\hat{y}[k] \cdot ((a_k^{pilot})^*)\}, \tag{57}$$

where  $a_k^{pilot}$  — transmitted pilot signal. The assessed phase is interpolated in the intermediate symbols between two subsequent pilots. In most recent papers, which discuss operation of formats that are more sensitive to phase distortions (16QAM, 64QAM etc.), it is suggested to use double-stage compensation circuits, where one stage is blind, and the other one uses pilots [72–74,76].

Note that this block is aimed at fighting phase distortions, regardless of the source of their occurrence, whether it is the laser emission band width or non-linear distortions in optical fiber. However, this block must not and may not fully compensate all phase distortions.

### 1.8. Full tract of DSP

Most papers on digital signal processing are dedicated to individual blocks that compensate specific distortion. Operability of these algorithms derives from an idealized signal, where often only the distortion that the block itself is responsible for is presented at the input to the block. Quality of distortion compensation is mainly tested only with numerical modeling. There are in general very few papers dedicated to review of DSP operation.



**Figure 8.** Circuit diagram of digital signal processing block operation. ADC — Analog-to-Digital Converter, AGG — Automatic Gain Control, CDC — Chromatic Dispersion Compensator, RCFO — Rough Carrier Frequency Offset, PCPR — Pilot Carrier Phase Recovery, BCPR — Blind Carrier Phase Recovery.

DSP circuits are presented in papers [52,57,77]. The papers [57,77] suggest the potential configuration of the circuit with a review of algorithms that may be used to fill in such circuits. No specific implementation of the operable circuit is provided in these papers. Using the proposed circuits in practice will bring the need to search for reliable algorithms for the specific implementation of the circuit. However, it is much more logical to search for the algorithms meeting the requirements, and to configure a full tract of DSP on their basis. Besides, we believe it is necessary to adapt the presented circuits for use with real signals. The paper [52] is one of few papers to present and research the implementation of the adaptive filter in the connection with the blocks it requires.

The objective for optimization of DSP block, including selection and optimization of algorithms of individual functional blocks, the sequence of their performance and optimization of their joint operation is non-trivial and does not have a single solution. Section 2 presents the original circuit of digital signal processing with description of the selected algorithms and their parameters.

## 2. Implementation

It should be noted once again that the objectives are not only to select and implement a set of algorithms, but also to agree these blocks within a single DSP circuit. As one may see, most presented algorithms decrease that distortion that they have to compensate, and the issue of their resistance to other, not yet compensated distortions, is usually not raised.

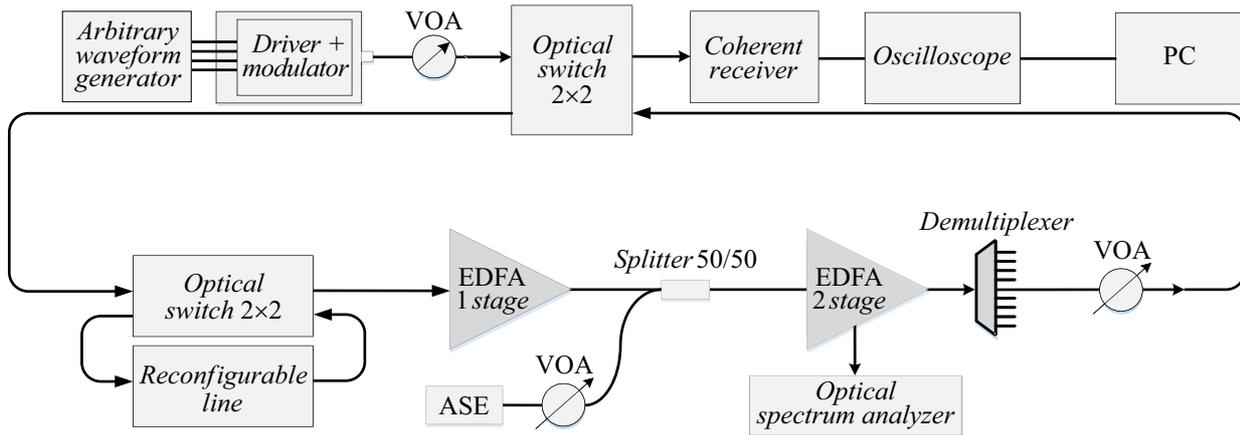
This paper suggested and researched a circuit shown in Fig. 8, which was developed with account of the joint stable operation of algorithms with experimental data. Modulation format — DP-QPSK.

For AGC, CAPS algorithm was selected and implemented, with the effective length of  $N_{AGC} = 1024$  sample averaging.

Chromatic dispersion was compensated in the frequency area with overlap of 50%. The length of the CD compensation block was selected as equal to  $N_{CDC} = 4096$  samples in order to compensate the dispersion in the line with length of 4000 km. If the distance is shorter, the filter length may be decreased.

At the next stage it was necessary to roughly estimate and compensate the carrier frequency offset, since the timing synchronization block operates worse under the conditions of the high carrier offset. Offset was estimated by searching for a peak in the spectrum of the fourth degree of the signal [59], the filter length was  $N_{RCFO} = 2048$  samples. Besides, for reliable and stable operation with high noise present, an accumulator was applied on the spectra (with a coefficient), as well as preliminary rough rotation of polarization with sorting at pitch of 10 degr.

Recovery of timing and cubic Farrows interpolation up to 2 samples per symbol was carried out using Gardner algorithm in parallel form [47] with parallelism depth  $N_{Ret} = 64$ . Gardner algorithm was selected since it was studied well in the temporary area and in the parallel form in the paper [47].



**Figure 9.** Circuit diagram of experimental unit to test operability of DSP algorithms.

An adaptive filter chosen was a  $4 \times 2$  circuit in the frequency area with CMA metric for tune-up. The metric was minimized by stochastic gradient descent with descent coefficient  $2^{-12}$ . Filter length was selected as equal to  $N_{AE} = 64$  samples. It should be noted that inside the adaptive filter, sampling decreased down to one sample per symbol, and overlap of 50% was used, i.e. the filter output contained 16 samples in every polarization.

Further the precise recovery of the carrier frequency took place, as well as rough recovery according to the algorithm [59]. The block length was  $N_{FCFO} = 128$ , each 32nd sample was used if time size of the block was 4096.

For frame synchronization, correlation was used with length of  $N_{FS} = 22$  symbols with service sequence FAW (frame alignment word) from [12].

The carrier phase was assessed in two stages: first using pilot symbols from [12] with the help of equation (57) with averaging by  $N_{PCPR} = 11$  assessments, and then using the algorithm with sorting of possible phase values in the quantity of  $N_{BCPR} = 40$  pcs [71].

It should be noted that all algorithms may be implemented in numbers with the fixed point and in the parallel form, i.e. in very-large-scale integrated circuits (VLSIC). All coherent communication systems at this moment use VLSIC exactly in part of signal processing to ensure maximum throughput capacity.

### 3. Operation results

The operation of the proposed circuit (Fig. 8) is tested on the experimental unit, the design of which is presented in Fig. 9. The signal is set by a generator, then a transmitter made of a driver and a DP-QPSK modulator, the signal is converted from an electric domain into an optical one, the signal is demodulated by a coherent receiver, and using an oscilloscope with the ADC capacity of 10 bit the data in the form of numbers is transmitted to a personal computer (PC), where they are processed by DSP, suggested in this paper.

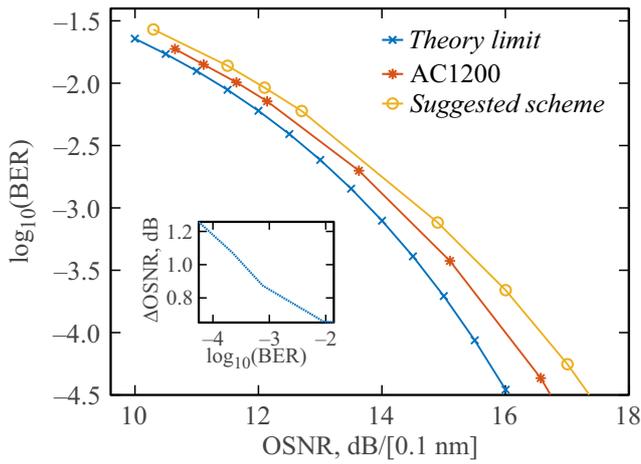
The experimental unit consists of several units: a reconfigurable line making it possible to change length from 0 to 4000 km, a PMD unit based on fiber maintaining the polarization state, setting a polarization-mode dispersion with value of 70 ps.

The unit also makes it possible to set the frequency offset between the laser on the transmitter and on the receiver, since the overtuned lasers are used according to ITLA standard [62]. Besides, previously the emission line width was measured for the used lasers, it turned out to be equal to 150 kHz on the transmitter and 150 kHz on the receiver.

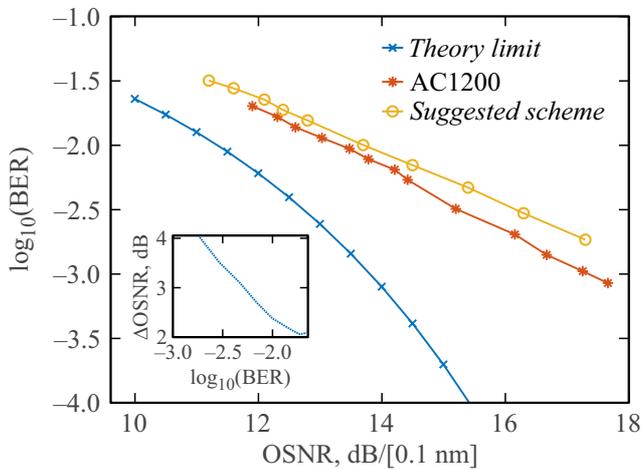
Several series of experiments were carried out with line length of 0, 2800, 4000 km, and also 2800 and 4000 km while the signal passes the PMD blocks. The main configurations of the experimental circuit are given below:

- shortest line (0 km);
- line with length of 2800 km:
  - without additions;
  - additional PMD 75 ps and offset of the local oscillator 5 GHz;
- line with length of 4000 km:
  - without additions;
  - additional PMD 75 ps and offset of the local oscillator 5 GHz.

To assess the quality of DSP operation, the dependence of the common logarithm of the bit error ratio  $BER$  on the optical signal to noise ratio  $OSNR$  is used, the noise capacity is converted into band 0.1 nm. The dependence  $BER(OSNR)$  is compared for the solution presented in this paper, the best commercially available transponder Acacia AC1200 and the theoretical limit, which presumes complete compensation of all distortions, except for the additive white Gaussian noise (electric noise of the transceiver is not taken into account). The theoretical limit was calculated as



**Figure 10.** Results of operation of the suggested algorithms in the line with length of 0 km.



**Figure 11.** Results of operation of the suggested algorithms in the line with length of 2800 km.

follows: from  $OSNR$  signal one may obtain  $SNR$  [78,79], from which it is possible to obtain  $BER$  [80]. The final equations to convert  $OSNR$  into  $BER$ , defining the theoretical limit of DSP operation for QPSK signal appear as

$$BER = 2Q\left(\sqrt{SNR}\right)\left(1 - \frac{1}{2}Q\left(\sqrt{SNR}\right)\right),$$

$$Q(x) = \frac{1}{2}\text{erfc}\left(x/\sqrt{2}\right),$$

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2/2) dt,$$

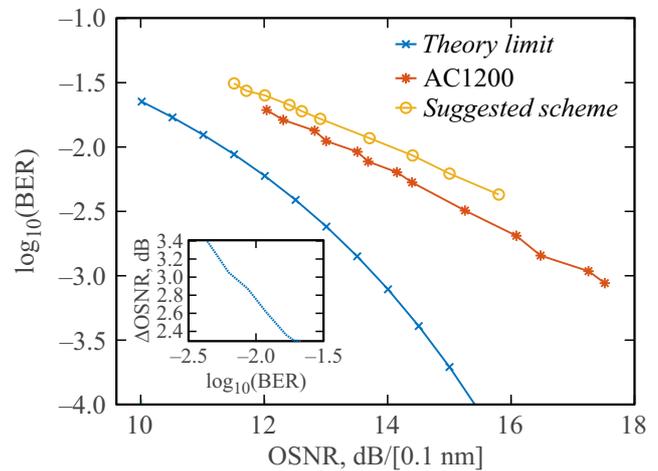
$$SNR = OSNR \frac{BW_{0.1}}{R_s}, \tag{58}$$

where  $SNR$  and  $OSNR$  are presented in linear units,  $BW_{0.1} = 12.5$  GHz — band width 0.1 nm in GHz on the wavelength of the carrier 1550 nm.

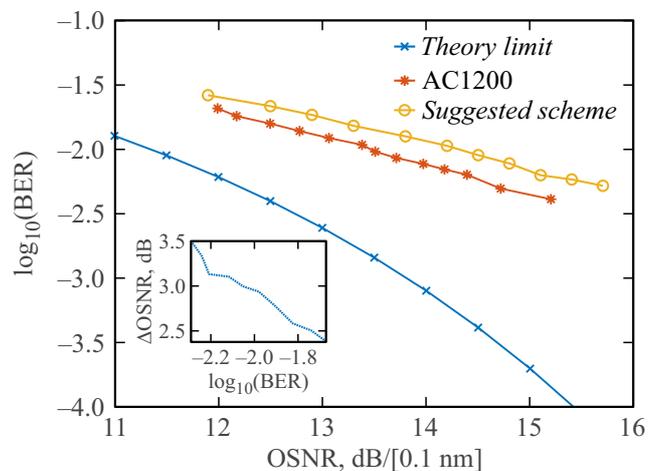
Dependences  $BER(OSNR)$  are given in Fig. 10–14, the chart attached presents the dependence of the penalty of the suggested circuit relative to the theoretical limit depending on level  $\log_{10}(BER)$ . The penalty is calculated as difference between  $OSNR$  on one level  $\log_{10}(BER)$ .

Results of comparison for level  $BER = 1.5\%$ , which is a median value of critical error ratio in operability of current FEC, used in the coherent optical communication, are presented in the table. For line with length of 0 km one may see that the results of the suggested circuit differ from the theoretical limit approximately by 0.7 dB, and from Acacia AC1200 - by 0.4 dB. For line with length of 2800 and 4000 km the penalty increase in relation to the theoretical limit is first of all dependent on the accumulated nonlinearity and amounts to the value of the order of 2–3 dB, however, there is also a penalty from PMD effects and deviation of the carrier frequency of the order of 0.3 dB.

The penalty 0.4–0.7 dB relative to AC1200 may arise for the following reasons:



**Figure 12.** Results of operation of the suggested algorithms in the line with length of 2800 km with PMD block and offset of the local oscillator.

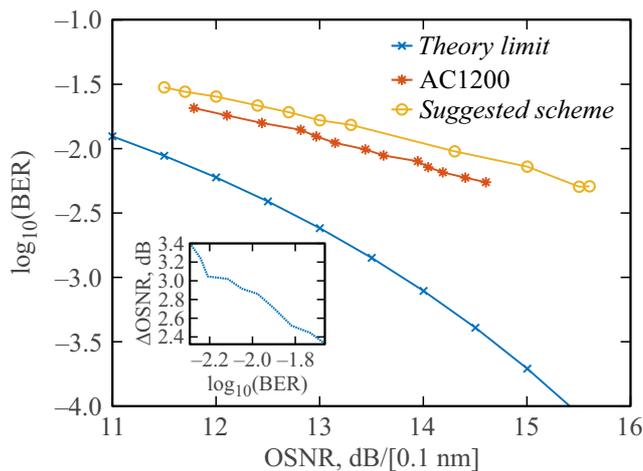


**Figure 13.** Results of operation of the suggested algorithms in the line with length of 4000 km.

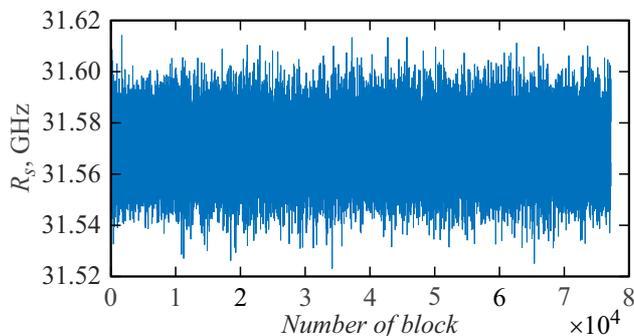
## Signal demodulation results

Number Figure	Length, km	PMD, ps	Carrier Deviation Frequency, GHz	Penalty to Theoretical Limit, dB	Penalty to AC1200, dB
10	0	0	< 0.4	0.7	0.4
11	2800	0	< 0.4	2.1	0.4
12	2800	75	5	2.4	0.6
13	4000	0	< 0.4	2.6	0.7
14	4000	75	5	2.6	0.7

Note. The penalty means difference by OSNR at level BER = 1.5% .



**Figure 14.** Results of operation of the suggested algorithms in the line with length of 4000 km with PMD block and offset of the local oscillator.



**Figure 15.** Stability of assessment of symbol speed  $R_s$  by parallel Gardner algorithm.

1) application of pre-distortions, making it possible to compensate losses and pulse response characteristic of the transmitter;

2) more stable timing synchronization;

3) partial compensation of non-linear effects;

4) more stable connections of the transmitter and receiver elements in AC1200 thanks to design in the form of a separate built-in block.

From the provided reasons, only more stable timing synchronization (Fig. 15) requires finalization of the presented circuit, the win of which will make around several tenths dB.

## Conclusion

The paper conducted the comparative analysis of the existing algorithms of linear DSP for coherent optical communication, following the results of which, the full-scale original circuit was suggested and implements for DP-QPSK signal. Operability of the suggested circuit is experimentally studied and makes it possible to transmit data at speed of 100G in real communication systems. It is shown that the suggested circuit successfully handles the compensation of physical distortions, arising in the communication line, to the level making it possible to detect the transmitted signal using the error ratio below the critical value. The penalty relative to the theoretical limit does not exceed 2.6 dB, and relative to the industry leader - 0.7dB in the line of up to 4000 km.

## References

- [1] VNI Complete Forecast Highlights (2020).
- [2] Cisco Annual Internet Report (2018-2023) White Paper (2020).
- [3] Fiber Optic Lines to Dominate Fixed Broadband Landscape in Asia-Pacific through 2026, forecasts GlobalData (2021).
- [4] T. Okoshi. J. Lightwave Technol., **5** (1), 44 (1987). DOI: 10.1109/JLT.1987.1075396
- [5] V. Konyshov, A. Leonov, O. Naniy, D. Starykh, V. Treshchikov, R. Ubaidullaev. Bull. Lebedev Phys. Institute, **50** (Suppl 4), S435 (2023). DOI: 10.3103/S1068335623160078
- [6] L. Wang, Y. Chen, X. Wang, C. Zhao, J. Chen, M. Tang. In *2021 Optical Fiber Communications Conference and Exhibition (OFC)* (IEEE, 2021), 1–3, DOI: 10.1364/OFC.2021.M5G.3
- [7] S. Tsukamoto, D.-S. Ly-Gagnon, K. Katoh, K. Kikuchi. In *Optical Fiber Communication Conference*, PDP29 (Optica Publishing Group, 2005), DOI: 10.1109/OFC.2005.193207
- [8] Y. Wakayama, T. Gerard, E. Sillekens, L. Galdino, D. Lavery, R.I. Killey, P. Bayvel. Opt. Express, **29** (12), 18743 (2021). DOI: 10.1364/OE.423361
- [9] X. Chen, S. Chandrasekhar, J. Cho, P. Winzer. Opt. Express, **27** (21), 29916 (2019). DOI: 10.1364/OE.27.029916

- [10] A. Ghazisaeidi, I.F. de Jauregui Ruiz, R. Rios-Müller, L. Schmalen, P. Tran, P. Brindel, A.C. Meseguer, Q. Hu, F. Buchali, G. Charlet, et al. *J. Lightwave Technol.*, **35** (7), 1291 (2017). DOI: 10.1109/JLT.2017.2657329
- [11] H. Sun, K.-T. Wu, K. Roberts. *Opt. express*, **16** (2), 873 (2008). DOI: 10.1364/OE.16.000873
- [12] *Open ROADM MSA 5.0 W-Port Digital Specification (100G–400G)* (2021).
- [13] M. Vucic, M. Butorac. In *2009 IEEE International Symposium on Circuits and Systems* (IEEE, 2009), 1032–1035, DOI: 10.1109/ISCAS.2009.5117935
- [14] T. Shan, T. Kailath. *IEEE Transactions on Circuits and Systems*, **35** (1), 122 (1988). DOI: 10.1109/31.1709
- [15] E. Tisserand, Y. Berville. *Electron. Lett.*, **52** (22), 1847 (2016). DOI: 10.1049/el.2016.1398
- [16] G.P. Agrawal. In *Nonlinear Science at the Dawn of the 21st Century* (Springer, 2000), 195–211, DOI: 10.1007/3-540-46629-0-9
- [17] H.M. Ozaktas, M.A. Kutay. In *2001 European Control Conference (ECC)* (IEEE, 2001), 1477–1483.
- [18] A. Yang, P. Guo, W. Wang, Y. Lu, Y. Qiao. *Optics Laser Technol.*, **111**, 447 (2019). DOI: 10.1016/j.optlastec.2018.10.021
- [19] F. Wu, A. Yang, P. Guo, Y. Qiao, L. Zhuang, S. Guo. *IEEE Photon. J.*, **10** (4), 1 (2018). DOI: 10.1109/JPHOT.2018.2859428
- [20] F. Wu, P. Guo, A. Yang, Y. Qiao. *IEEE Access*, **7**, 139388 (2019). DOI: 10.1109/ACCESS.2019.2944026
- [21] H. Zhou, B. Li, M. Tang, K. Zhong, Z. Feng, J. Cheng, A.P.T. Lau, C. Lu, S. Fu, P.P. Shum, et al. *J. Lightwave Technol.*, **34** (10), 2371 (2016). DOI: 10.1109/JLT.2016.2538467
- [22] D. Wang, C. Lu, A.P.T. Lau, S. He. *IEEE Photon. Technol. Lett.*, **23** (14), 1016 (2011). DOI: 10.1109/LPT.2011.2151280
- [23] C. Xie. *IEEE Photon. Technol. Lett.*, **25** (10), 992 (2013). DOI: 10.1109/LPT.2013.2257729
- [24] J.C. Diniz, S.M. Ranzini, V. Ribeiro, E. Magalhaes, E.S. Rosa, V. Parahyba, L. Franz, E. Ferreira, J.C. Oliveira. In *Optical Fiber Communication Conference, OTh3C–6* (Optica Publishing Group, 2013), DOI: 10.1364/OFC.2013.OTh3C.6
- [25] D. Wang, C. Lu, A.P.T. Lau, P.K.A. Wai, S. He. *IEEE Photon. Technol. Lett.*, **25** (10), 985 (2013). DOI: 10.1109/LPT.2013.2257721
- [26] C. Malouin, P. Thomas, B. Zhang, J. O’Neil, T. Schmidt. In *Signal Processing in Photonic Communications, SpTh2B–4* (Optica Publishing Group, 2012), DOI: 10.1364/SPPCOM.2012.SpTh2B.4
- [27] Q. Sui, A.P.T. Lau, C. Lu. *J. Lightwave Technol.*, **31** (2), 306 (2012). DOI: 10.1109/JLT.2012.2231400
- [28] F.N. Hauske, M. Kuschnerov, B. Spinnler, B. Lankl. *J. Lightwave Technol.*, **27** (16), 3623 (2009). DOI: 10.1109/JLT.2009.2024960
- [29] C.C. Do, C. Zhu, A.V. Tran, S. Chen, T. Anderson, D. Hewitt, E. Skafidas. In *OFC/NFOEC* (IEEE, 2012), 1–3, DOI: 10.1364/OFC.2012.OW4G.1
- [30] M. Golay. *IRE Transactions on Information Theory*, **7** (2), 82 (1961). DOI: 10.1109/TIT.1961.1057620
- [31] T. Xu, G. Jacobsen, S. Popov, J. Li, E. Vanin, K. Wang, A.T. Friberg, Y. Zhang. *Opt. Express*, **18** (15), 16243 (2010). DOI: 10.1364/OE.18.016243
- [32] S.J. Savory. *Opt. Express*, **16** (2), 804 (2008). DOI: 10.1364/OE.16.000804
- [33] V.A. Kotelnikov. *Phys. Usp.*, **49** (7), 736–744, (2006). DOI: 10.1070/PU2006v049n07ABEH0006160
- [34] R.D. Strum, D.E. Kirk. *First Principles of Discrete Systems and Digital Signal Processing* (Addison-Wesley Longman Publishing Co., Inc., 1988)
- [35] R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, Y. Miyamoto. *J. Lightwave Technol.*, **27** (16), 3721 (2009). DOI: 10.1109/JLT.2009.2024091
- [36] K. Mueller, M. Muller. *IEEE Transactions on Commun.*, **24** (5), 516 (1976). DOI: 10.1109/TCOM.1976.1093326
- [37] F. Gardner. *IEEE Transactions on Commun.*, **34** (5), 423 (1986). DOI: 10.1109/TCOM.1986.1096561
- [38] M.G. Floyd. *IEEE Transactions on Commun.*, **41** (3), 501 (1993). DOI: 10.1109/26.221081
- [39] M. Oerder, H. Meyr. *IEEE Transactions on Commun.*, **36** (5), 605 (1988). DOI: 10.1109/26.1476
- [40] S.J. Lee. *IEEE Commun. Lett.*, **6** (5), 205 (2002). DOI: 10.1109/4234.1001665
- [41] D. Godard. *IEEE Transactions on Commun.*, **26** (5), 517 (1978). DOI: 10.1109/TCOM.1978.1094107
- [42] L. Huang, D. Wang, A.P.T. Lau, C. Lu, S. He. *Opt. Express*, **22** (6), 6749 (2014). DOI: 10.1364/OE.22.006749
- [43] F.M. Gardner. *IEEE Transactions on Commun.*, **41** (3), 501 (1993). DOI: 10.1109/26.221081
- [44] L. Erup, F.M. Gardner, R.A. Harris. *IEEE Transactions on Commun.*, **41** (6), 998 (1993). DOI: 10.1109/26.231921
- [45] X. Zhou, X. Chen, W. Zhou, Y. Fan, H. Zhu, Z. Li. *J. Optical Commun. Networking*, **2** (11), 984 (2010). DOI: 10.1364/JOCN.2.000984
- [46] X. Zhou. *IEEE Signal Processing Magazine*, **31** (2), 35 (2014). DOI: 10.1109/MSP.2013.2281071
- [47] X. Zhou, X. Chen. *Opt. Express*, **19** (10), 9282 (2011). DOI: 10.1364/OE.19.009282
- [48] C.W. Farrow. In *1988, IEEE International Symposium on Circuits and Systems* (IEEE, 1988), 2641–2645, DOI: 10.1109/ISCAS.1988.15483
- [49] J. Gordon, H. Kogelnik. *Proceed. National Academy of Sci.*, **97** (9), 4541 (2000). DOI: 10.1073/pnas.97.9.4541
- [50] R. Rios-Müller, J. Renaudier, G. Charlet. *J. Lightwave Technol.*, **33** (7), 1315 (2015). DOI: 10.1109/JLT.2014.2377582
- [51] D. Godard. *IEEE Transactions on Commun.*, **28** (11), 1867 (1980). DOI: 10.1109/TCOM.1980.1094608
- [52] R. Babae, S.O. Gharan, M. Bouchard. *J. Lightwave Technol.*, (2023). DOI: 10.1109/JLT.2023.3253383
- [53] M.S. Faruk, K. Kikuchi. *Opt. Express*, **19** (13), 12789 (2011). DOI: 10.1364/OE.19.012789
- [54] M.S. Faruk, K. Kikuchi. In *2014 The European Conference on Optical Communication (ECOC)* (IEEE, 2014), 1–3, DOI: 10.1109/ECOC.2014.6964149
- [55] K. Kikuchi. *Opt. Express*, **19** (10), 9868 (2011). DOI: 10.1364/OE.19.009868
- [56] L. Liu, Z. Tao, W. Yan, S. Oda, T. Hoshida, J.C. Rasmussen. In *2009 Conference on Optical Fiber Communication* (IEEE, 2009), 1–3, DOI: 10.1364/OFC.2009.OMT2
- [57] S.J. Savory. *IEEE J. of Selected Topics in Quantum Electronics*, **16** (5), 1164 (2010). DOI: 10.1109/JSTQE.2010.2044751

- [58] H.L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory* (John Wiley & Sons, 2004)
- [59] M. Morelli, U. Mengali. *European Transactions on Telecommunications*, **9** (2), 103 (1998). DOI: 10.1002/ett.4460090203
- [60] N. Noels, H. Steendam, M. Moeneclaey, H. Bruneel. *IEEE Transactions on Signal Processing*, **53** (12), 4578 (2005). DOI: 10.1109/TSP.2005.859318
- [61] F. Rice. *IEEE Transactions on Commun.*, **54** (2), 221 (2006). DOI: 10.1109/TCOMM.2005.863782
- [62] MSA. *Optical Internetworking Forum, OIF* (July 13th, 2015)
- [63] R. Frank, S. Zedoff, R. Heilmiller. *IRE Transactions on Information Theory*, **8** (6), 381 (1962). DOI: 10.1109/TIT.1962.1057786
- [64] F. Pittalà, F.N. Hauske, Y. Ye, N.G. Gonzalez, I.T. Monroy. In *OFC/NFOEC* (IEEE, 2012), 1–3, DOI: 10.1364/OFC.2012.OM2H.4
- [65] O. Riznyk, Y. Kynash, B. Balych, R. Vynnychuk, I. Kret, H. Kutucu. In *CEUR Workshop Proceedings* (2019), 162–171, DOI: 10.1109/PICST47496.2019.9061375
- [66] T.M. Schmidl, D.C. Cox. *IEEE Transactions on Commun.*, **45** (12), 1613 (1997). DOI: 10.1109/26.650240
- [67] Q. Zhuge, M. Morsy-Osman, X. Xu, M. Chagnon, M. Qiu, D.V. Plant. *J. Lightwave Technol.*, **31** (15), 2621 (2013). DOI: 10.1109/JLT.2013.2271634
- [68] Z. Gao, C. Zhang, Z. Wang. *IEEE Transactions on Broadcasting*, **61** (1), 98 (2015). DOI: 10.1109/TBC.2014.2376134
- [69] A.J. Viterbi, A.M. Viterbi. *IEEE Transactions on Information Theory*, **29** (4), 543 (1983). DOI: 10.1109/TIT.1983.1056713
- [70] E. Ip, J.M. Kahn. *J. Lightwave Technol.*, **25** (9), 2675 (2007). DOI: 10.1109/JLT.2007.902118
- [71] T. Pfau, S. Hoffmann, R. Noé. *J. Lightwave Technol.*, **27** (8), 989 (2009). DOI: 10.1109/JLT.2008.2010511
- [72] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, M. Bertolini, G. Gavioli. *IEEE Photon. Technol. Lett.*, **24** (9), 739 (2012). DOI: 10.1109/LPT.2012.2187439
- [73] H. Cheng, Y. Li, M. Yu, J. Zang, J. Wu, J. Lin. In *Optical Fiber Communication Conference, Th4D-1* (Optica Publishing Group, 2014), DOI: 10.1364/OFC.2014.M2I.2
- [74] F. Guiomar, M. Neves, A. Lorences-Riesgo, C. Martins, S. Mumtaz, Y. Frignac, G. Charlet, P. Monteiro. In *2023 Optical Fiber Communications Conference and Exhibition (OFC)* (IEEE, 2023), 1–3, DOI: 10.1364/OFC.2023.W3E.1
- [75] A. Spalvieri, L. Barletta. *IEEE Transactions on Commun.*, **59** (7), 1966 (2011). DOI: 10.1109/TCOMM.2011.051311.100047
- [76] P. Gianni, G. Corral-Briones, C. Rodriguez, M.R. Hueda. *IEEE Photon. Technol. Lett.*, **25** (5), 442 (2013). DOI: 10.1109/LPT.2013.2241050
- [77] D.A. Morero, M.A. Castrillon, A. Aguirre, M.R. Hueda, O.E. Agazzi. *J. Lightwave Technol.*, **34** (1), 121 (2016). DOI: 10.1109/JLT.2015.2470114
- [78] E. Torrenco, R. Cigliutti, G. Bosco, A. Carena, V. Curri, P. Poggiolini, A. Nespola, D. Zeolla, F. Forghieri. *Opt. Express*, **19** (26), B790 (2011). DOI: 10.1364/OE.19.00B790
- [79] F. Vacondio, O. Rival, C. Simonneau, E. Grellier, A. Bononi, L. Lorcy, J.-C. Antona, S. Bigo. *Opt. Express*, **20** (2), 1022 (2012). DOI: 10.1364/OE.20.001022
- [80] J.G. Proakis, M. Salehi. *Digital Commun.* (McGraw-Hill, 2008)

Translated by M. Verenikina