Global Estimation and Compensation of Linear Effects in Coherent Optical Systems based on Nonlinear Least Squares

Global Estimation and Compensation of Linear Effects in Coherent Optical Systems based on Nonlinear Least Squares

Alexandru Frunza, *Student Member, IEEE,* Vincent Choqueuse, *Member, IEEE,* Pascal Morel, *Member, IEEE,*, Stéphane Azou, *Senior Member, IEEE,*

Abstract—This paper proposes a new estimation and compensation approach to mitigate several linear and widely linear effects in coherent optical systems using digital signal processing (DSP) algorithms. Compared to most of the available strategies that employ local estimation and/or compensation algorithms, this approach performs a global impairments estimation and compensation based on Nonlinear Least Squares. The proposed method estimates and compensates for the chromatic dispersion (CD), carrier frequency offset (CFO), in-phase/quadrature (IQ) imbalance, and laser phase noise (PN) in two steps. Firstly, it estimates the quasi-static parameters related to the CD, CFO, and both transmitter and receiver IQ imbalance. Secondly, it estimates both transmitter and receiver lasers' phases and compensates for all the imperfections by using a Zero-Forcing (ZF) equalizer. Simulations show the effectiveness of the approach in terms of statistical performance and computational time. The estimation performance is assessed by computing the Cramér Rao Lower Bound (CRLB), while the detection performance is compared to a modified Clairvoyant equalizer.

Index Terms—Coherent optical systems, global estimation and compensation, chromatic dispersion, carrier frequency offset, IQ imbalance, laser phase noise, Nonlinear Least Squares, optimization.

I. INTRODUCTION

The rapid growth of the Internet traffic and the emerging technologies as 5G and the Internet of Things [1] lead to an increasing demand for high data rate communications. Coherent optical fiber transmission systems are thought to play a central role in meeting this challenge due to some key features, including huge available bandwidth and the support of advanced modulation formats. However, for systems with high order modulation formats, the performance can be drastically impaired by hardware imperfections and is more sensitive to optical channel induced effects. Among the most important impairments are the in-phase/quadrature (IQ) amplitude and phase imbalance [2], the IQ time skew [3], the chromatic dispersion (CD) [4], the carrier frequency offset (CFO) [5], the laser phase noise (PN) [6], and the fiber nonlinearity [7]. To mitigate these imperfections, different compensation techniques have been proposed. Although many studies focus on the compensation of nonlinear effects, a large part of the optical communication chain's impairments can be modeled,

A. Frunza, V. Choqueuse, P. Morel, and S. Azou are with the Lab-STICC, UMR CNRS 6285, ENIB, Technopole Brest-Iroise, CS 73862, 29238 Brest Cedex 3, France (e-mail:frunza@enib.fr, choqueuse@enib.fr, morel@enib.fr, azou@enib.fr). A. Frunza is also with the C.E.C.T.I, Military Technical Academy, 050141 Bucharest, Romania.

as a first approximation, by linear or widely linear effects. The linear effects can include the IQ time skew, CD, CFO, and PN, while the IQ imbalance can be described as a widely linear effect. Impairments compensation can be performed using analog or digital techniques. The digital signal processing (DSP) methods have no loss in performance compared to the analog techniques and have the advantage of being more flexible. Furthermore, the recent advances in the development of high-speed integrated circuits make DSP compensation feasible at high data rates [8]–[10].

Recent works have proposed multiple full DSP or hybrid methods to mitigate the impact of linear and widely linear imperfections. These methods mainly focus on a few numbers of imperfections of the optical chain, employing local compensation techniques. Regarding the IQ imbalance, several approaches have been proposed for its compensation, including Gram-Schmidt orthogonalization procedure [11], ellipse correction [12], adaptive filtering [13]–[15], Kalman filtering [16], blind adaptive source separation [17], and machine learning-based [18]. Different techniques have also been proposed for IQ time skew estimation and compensation [19]–[22]. For the CD, which is traditionally compensated for by using optical fibers with opposite dispersion [23], digital filtering algorithms were employed in [4], [24]–[26]. Various methods for CFO and PN compensation were recently proposed in [27]–[32], and [33]–[37], respectively. Using a local compensation technique can be problematic. The performance of local compensation algorithms can be significantly reduced in the presence of other imperfections. Moreover, the integration of multiple DSP algorithms needed to compensate for all the impairments can be challenging. Furthermore, some of the compensation methods referenced above are optimized for a specific modulation format, and by this, their applicability can be limited to particular scenarios.

To alleviate these problems, this work proposes a global technique that can jointly estimate and compensate for multiple impairments, including the IQ imbalance, CD, CFO, and PN. The IQ time skew could be compensated before the proposed algorithm as in [20]. Our method is essentially designed for single carrier (SC) coherent optical systems and is modulation-format independent. The parameters describing the impairments are separated into quasi-static (CD, CFO, IQ imbalance) and time-variant (PN) parameters. The estimation is based on Nonlinear Least Squares. Firstly, the quasi-static parameters are globally estimated using a preamble. During
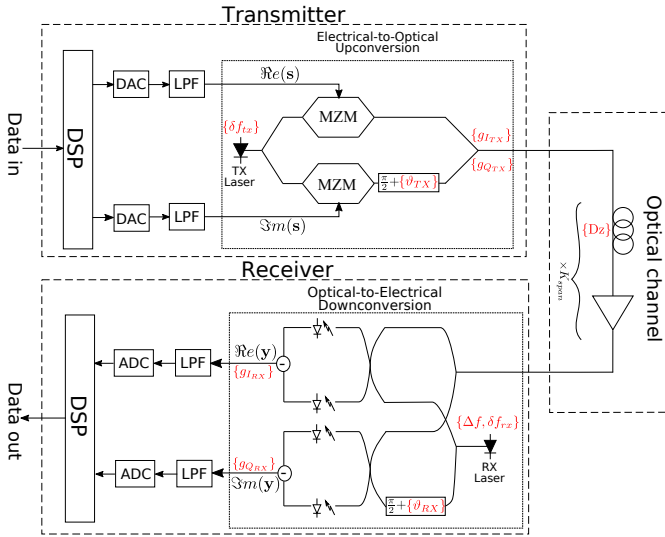
Fig. 1: Block diagram of a coherent optical system under the impact of the laser phase noise (PN), IQ imbalance, chromatic dispersion (CD), and carrier frequency offset (CFO)

this estimation step, the impact of PN is also considered. Secondly, by using the estimated quasi-static parameters, as the PN is a time-variant effect, the evolution of the laser phase is tracked over the whole data sequence using pilots. The compensation is based on the Zero-Forcing (ZF) equalizer and detects the transmitted symbols by inverting the global channel impact. As compared to other techniques, our proposed method has the advantage of jointly tracking both transmitter and receiver lasers' phases.

The proposed technique has a great potential for a wide range of applications, as it can mitigate several imperfections, shows good statistical performances, is modulation-format independent and has a limited computational time. The proposed method is well suited for metro systems and data center fiber technology systems, as it is operated in the linear regime and can compensate for multiple impairments. Moreover, the proposed algorithm can be used in future flexible optical systems based on its modulation-format independence.

The paper is organized as follows. Section II describes the signal model under the impact of the CD, CFO, transmitter and receiver IQ imbalance and PN. The proposed estimation and compensation algorithms are derived in Section III. The effectiveness of the method is validated by numerical simulation in section IV, while section V concludes the paper.

II. SIGNAL MODEL

In this paper, we focus on a coherent optical communication system under the impact of PN and IQ imbalance, both on the transmitter and receiver sides, CFO between the transmitter and receiver lasers, and the CD induced by the optical channel. The block diagram of the proposed communication system is presented in figure 1. The modeling of various impairments is discussed in II-A, and the general signal model of the communication is presented in subsection II-B.

A. Impairments model

A linear effect can be described by $\mathbf{y} = \mathbf{H}(\boldsymbol{\alpha})\mathbf{s}$, where $\mathbf{y}$ is the signal at the output of the system, $\mathbf{H}(\boldsymbol{\alpha})$ is the transfer matrix that depends nonlinearly on the vector of parameters $\boldsymbol{\alpha}$ to be estimated, and $\mathbf{s}$ is the input signal. A widely linear effect can be described by $\tilde{\mathbf{y}} = \tilde{\mathbf{M}}(\boldsymbol{\alpha})\tilde{\mathbf{s}}$, where the tilde denotes the augmented vectors or matrices containing the real and imaginary parts of the original vectors, and $\tilde{\mathbf{M}}(\boldsymbol{\alpha})$ is an augmented matrix related to the transfer matrix. In our work, the system equation can be written as $\mathbf{y} = \mathbf{A}(\boldsymbol{\beta}, \boldsymbol{\phi})\boldsymbol{\theta}$, where $\mathbf{A}(\boldsymbol{\beta}, \boldsymbol{\phi})$ is a matrix containing the transmitted signal that depends on the nonlinear parameters $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$ is a column vector containing the linear parameters describing the IQ imbalance. The nonlinear parameters $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$ are related to CD and CFO, and PN, respectively.

The impairments modelization is done considering the impact of the imperfections on a generic input signal denoted as:

$$\mathbf{x}_{in} = \begin{bmatrix} x_{in}[0] & x_{in}[1] & \dots & x_{in}[N-1] \end{bmatrix}^T. \tag{1}$$

where N is the total number of signal samples, and $(.)^T$ denotes the transpose operation. The impaired signal is denoted by $\mathbf{x}_{out}$. The parameters related to the imperfections are divided into two categories regarding the time evolution of their values. The first category refers to the quasi-static parameters, the parameters whose values have a very slow variation in time. This category includes the CD, CFO, and IQ imbalance parameters. The second category refers to the time-variant parameters, the parameters whose values vary relatively fast in time. This category includes PN.

1) Laser phase noise: One of the most critical impairments that impact coherent optical systems is the PN induced by both the transmitter and receiver lasers. The laser PN is a time-variant effect referring to the optical source frequency fluctuation. It is usually characterized by the laser linewidth, which is equal to 0 Hz in the ideal case. The PN is usually modeled as a Wiener process as follows [27], [33]:

$$\phi_k = \sum_{i=-\infty}^{k} f_i. \tag{2}$$

where the f_i's are independent and identically distributed random Gaussian variables with zero mean and variance $\sigma_f^2 = 2\pi \left(\frac{\delta f}{f_s} \right)$, with δf being the linewidth of the laser and f_s the sampling frequency. In high baud-rate optical communication systems, the carrier phase changes slowly compared to the signal phase, and it can be assumed constant over several K consecutive symbols [9], [38]. Under this assumption, the output samples can be modeled according to the input samples as:

$$\mathbf{x}_{out} = \boldsymbol{\Phi}(\boldsymbol{\phi})\mathbf{x}_{in}. \tag{3}$$

where $\boldsymbol{\Phi}(\boldsymbol{\phi})$ is an $N \times N$ diagonal matrix which is defined as:

$$\boldsymbol{\Phi}(\boldsymbol{\phi}) = \text{diag}(e^{j\phi_0}, e^{j\phi_1}, \dots, e^{j\phi_{N/K-1}}) \otimes \mathbf{I}_K. \tag{4}$$

with $\boldsymbol{\phi} = \begin{bmatrix} \phi_0 & \phi_1 & \dots & \phi_{N/K-1} \end{bmatrix}^T$, and $\mathbf{I}_K$ is a $K \times K$ identity matrix. The $\text{diag}(.)$ notation denotes a diagonal matrix

having on its main diagonal the values from its argument, and $\otimes$ corresponds to the Kronecker product.

2) IQ imbalance: IQ amplitude and phase imbalances are quasi-static impairments limiting the system performance that arise both on the transmitter and receiver side. Amplitude imbalance g refers to the amplitude mismatch between the I and Q branches of the optical modulator. The phase imbalance ϑ refers to the phase deviation from the ideal $90°$ between the branches. Starting from these, the IQ distortion parameters can be designed in an analytical convenable manner as in [39], [40]:

$$\mu = \cos\left(\frac{\vartheta}{2}\right) + jg\sin\left(\frac{\vartheta}{2}\right), \tag{5a}$$

$$\nu = g\cos\left(\frac{\vartheta}{2}\right) - j\sin\left(\frac{\vartheta}{2}\right). \tag{5b}$$

where $(\mu, \nu) \in \mathbb{C}^2$. After IQ imbalance, the output samples can be expressed according to the input samples as a widely linear transformation as follows [41], [42]:

$$\mathbf{x}_{out} = \mu\mathbf{x}_{in} + \nu\mathbf{x}_{in}^*. \tag{6}$$

where $(.)^*$ denotes the complex conjugate operation.

3) Chromatic dispersion: The dispersion in time of the fiber propagating signals results in a frequency dependency of the group velocity. This impairment is called CD. The CD is a quasi-static impairment that can limit the transmission distance and data rate and can be modeled by a filter whose transfer function is given by [24], [43]:

$$G_{\text{Dz}}(\omega) = e^{-j\frac{\text{Dz}\lambda^2}{4\pi c}\omega^2}. \tag{7}$$

with Dz being the accumulated CD coefficient, λ the wavelength, c the speed of the light, and ω the angular frequency with respect to f_s. The output samples can be expressed with respect to the input samples as:

$$\mathbf{x}_{out} = \mathbf{W}^H\mathbf{D_1}(\text{Dz})\mathbf{W}\mathbf{x}_{in}. \tag{8}$$

where $(.)^H$ denotes the conjugate transpose (Hermitian). In this model, the effect of the CD is computed in the frequency domain. Specifically, the DFT of the signal is first computed by multiplying the input signal $\mathbf{x}_{in}$ by the DFT matrix $\mathbf{W}$, then the DFT is multiplied with the frequency response of the chromatic dispersion $\mathbf{D_1}(\text{Dz})$, and finally, the time-domain signal is extracted from the multiplication with the inverse DFT matrix $\mathbf{W}^H$. Mathematically, the matrices $\mathbf{W}$ and $\mathbf{D_1}(\text{Dz})$ are given as follows

- $\mathbf{W}$ is an $N \times N$ Vandermonde matrix that corresponds to the DFT matrix and can be defined as:

$$\mathbf{W} = \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & z & & z^{N-1} \\ 1 & z^2 & & z^{2(N-1)} \\ \vdots & \vdots & & \vdots \\ 1 & z^{N-1} & \dots & z^{(N-1)^2} \end{bmatrix} \tag{9}$$

where $z = e^{-2j\pi/N}$. This matrix computes the DFT in the angular frequency range $\left[0, 2\pi\frac{(N-1)f_s}{N}\right]$ in rad/s.

- $\mathbf{D_1}(\text{Dz})$ is an $N \times N$ diagonal matrix that contains the effect of the CD. The n-th diagonal element of $\mathbf{D_1}(\text{Dz})$ can be expressed by:

$$[D_1(\text{Dz})]_{nn} = \begin{cases} G_{\text{Dz}}(2\pi nf_s/N) & \text{if } n < \frac{N}{2} \\ G_{\text{Dz}}(2\pi(n-N)f_s/N) & \text{if } n \geq \frac{N}{2} \end{cases}. \tag{10}$$

4) Carrier frequency offset: The free-running transmitter and receiver lasers in a coherent optical system are not frequency locked and, generally, a CFO cannot be neglected. The CFO can be described as a quasi-static impairment that decreases the system performance and should be estimated and compensated for on the receiver side. The CFO impact can be modeled as a phase increment as follows: [5], [44]

$$\phi_{k+1} = 2\pi\left(\frac{\Delta f}{f_s}\right) + \phi_k. \tag{11}$$

where Δf is the frequency offset. In this context, the relation between the input samples and output samples can be modeled into a matrix form as:

$$\mathbf{x}_{out} = \mathbf{D_2}(\Delta f)\mathbf{x}_{in}. \tag{12}$$

where $\mathbf{D_2}(\Delta f)$ is an $N \times N$ diagonal matrix whose n-th diagonal element is given by:

$$[D_2(\Delta f)]_{nn} = e^{j\frac{2\pi n\Delta f}{f_s}}. \tag{13}$$

B. System model

Let us denote the transmitted signal as follows:

$$\mathbf{s} = \begin{bmatrix} s[0] & s[1] & \dots & s[N-1] \end{bmatrix}^T. \tag{14}$$

During the electrical-to-optical conversion, the signal is impaired by the transmitter laser PN and can experience IQ imbalance. Therefore, the transmitted block of data can be expressed in two steps as:

$$\mathbf{u} = \mathbf{\Phi}(\boldsymbol{\phi_{tx}})\mathbf{s}, \tag{15}$$

$$\mathbf{z} = \mu_{tx}\mathbf{u} + \nu_{tx}\mathbf{u}^*. \tag{16}$$

Next, the combined effect of CD and CFO is accumulated and can be modeled as:

$$\mathbf{r} = \mathbf{H}(\boldsymbol{\beta})\mathbf{z}. \tag{17}$$

where $\mathbf{H}(\boldsymbol{\beta})$ is an $N \times N$ matrix that depends on the effects of the CD and CFO, and $\boldsymbol{\beta} = \begin{bmatrix} \text{Dz} & \Delta f \end{bmatrix}$. The matrix $\mathbf{H}(\boldsymbol{\beta})$ can be decomposed as:

$$\mathbf{H}(\boldsymbol{\beta}) = \mathbf{D_2}(\Delta f)\mathbf{W}^H\mathbf{D_1}(\text{Dz})\mathbf{W}. \tag{18}$$

Finally, by considering the presence of receiver front-end IQ imbalance and laser PN, the received signal can be expressed in two steps as:

$$\mathbf{v} = \mu_{rx}\mathbf{r} + \nu_{rx}\mathbf{r}^*, \tag{19}$$

$$\mathbf{y} = \mathbf{\Phi}(\boldsymbol{\phi_{rx}})\mathbf{v}. \tag{20}$$

Bringing all together, the received signal can be rewritten as follows:

$$\mathbf{y} = \mathbf{A}(\boldsymbol{\beta}, \boldsymbol{\phi})\boldsymbol{\theta} + \mathbf{b}. \tag{21}$$
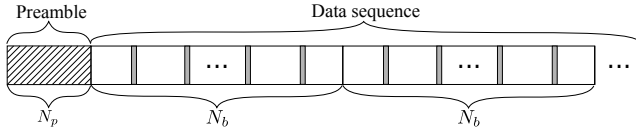
Fig. 2: The frame structure

where

- $\mathbf{A}(\boldsymbol{\beta}, \boldsymbol{\phi})$ is an $N \times 4$ matrix that contains the effect of the CD, CFO, and PN which is defined as:

$$\mathbf{A}(\boldsymbol{\beta}, \boldsymbol{\phi}) = \begin{bmatrix} \boldsymbol{\Phi}(\boldsymbol{\phi_{rx}})\mathbf{H}(\boldsymbol{\beta})\boldsymbol{\Phi}(\boldsymbol{\phi_{tx}})\mathbf{s} \\ \boldsymbol{\Phi}(\boldsymbol{\phi_{rx}})\mathbf{H}(\boldsymbol{\beta})\boldsymbol{\Phi}^*(\boldsymbol{\phi_{tx}})\mathbf{s}^* \\ \boldsymbol{\Phi}(\boldsymbol{\phi_{rx}})\mathbf{H}^*(\boldsymbol{\beta})\boldsymbol{\Phi}^*(\boldsymbol{\phi_{tx}})\mathbf{s}^* \\ \boldsymbol{\Phi}(\boldsymbol{\phi_{rx}})\mathbf{H}^*(\boldsymbol{\beta})\boldsymbol{\Phi}(\boldsymbol{\phi_{tx}})\mathbf{s} \end{bmatrix}^T . \quad (22)$$

  with $\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi}_{tx}^T & \boldsymbol{\phi}_{rx}^T \end{bmatrix}^T$ denoting the transmitter and receiver lasers' phases,

- $\boldsymbol{\theta}$ is a column vector that contains the IQ imbalance parameters and can be expressed as:

$$\boldsymbol{\theta} = \begin{bmatrix} \mu_{rx}\mu_{tx} \\ \mu_{rx}\nu_{tx} \\ \nu_{rx}\mu_{tx}^* \\ \nu_{rx}\nu_{tx}^* \end{bmatrix} . \quad (23)$$

- $\mathbf{b}$ is a column vector containing the noise impact.

The vectors $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$ contain the nonlinear parameters, and the vector $\boldsymbol{\theta}$ contains the linear parameters of the model. In this paper, the equation (21) will serve as a basis for the proposed estimation and compensation algorithms.

III. ESTIMATION AND COMPENSATION ALGORITHMS

In this section, the estimation and compensation algorithms are presented. It is assumed that the signal has the particular structure described in figure 2. The proposed method uses a number of N_p symbols as a preamble to estimate the system's quasi-static parameters, and L pilots from the N_b symbols to track the lasers' phases. On the receiver side, the preamble and pilots are assumed to be known. The synchronization could be performed based on the signal cross-correlation and interpolation techniques by using a dedicated preamble [45], [46].

A. Estimation algorithm

To estimate the quasi-static system parameters, we use a random training sequence of N_p symbols. As the training sequence is generally short, we consider the carrier phase to be constant throughout the entirety of it. Assuming this, the matrices $\boldsymbol{\Phi}(\boldsymbol{\phi}_{tx})$ and $\boldsymbol{\Phi}(\boldsymbol{\phi}_{rx})$ reduce to scalar matrices as follows:

$$\boldsymbol{\Phi}(\boldsymbol{\phi}_{tx}) = e^{j\phi_{tx,0}}\mathbf{I}_{N_p}, \quad (24a)$$
$$\boldsymbol{\Phi}(\boldsymbol{\phi}_{rx}) = e^{j\phi_{rx,0}}\mathbf{I}_{N_p}. \quad (24b)$$

In this context, the impact of the PN can be accumulated on the IQ phase imbalance and the received signal can be rewritten as:

$$\mathbf{y} = \mathbf{A}(\boldsymbol{\beta})\boldsymbol{\theta}_{\phi} + \mathbf{b}. \quad (25)$$

where

- $\mathbf{A}(\boldsymbol{\beta}) = \mathbf{A}(\boldsymbol{\beta}, 0)$ is an $N_p \times 4$ matrix that contains the effects of the CD and CFO, and is defined as:

$$\mathbf{A}(\boldsymbol{\beta}) = \begin{bmatrix} \mathbf{H}(\boldsymbol{\beta})\mathbf{s} & \mathbf{H}(\boldsymbol{\beta})\mathbf{s}^* & \mathbf{H}^*(\boldsymbol{\beta})\mathbf{s}^* & \mathbf{H}^*(\boldsymbol{\beta})\mathbf{s} \end{bmatrix} . \quad (26)$$

- $\boldsymbol{\theta}_{\phi}$ is a column vector that contains the IQ imbalance parameters mixed with the lasers' PNs, and can be expressed as:

$$\boldsymbol{\theta}_{\phi} = \mathbf{D}_{\phi}\boldsymbol{\theta}. \quad (27)$$

  with $\mathbf{D}_{\phi}$ being a diagonal matrix defined as:

$$\mathbf{D}_{\phi} = e^{j\phi_{rx,0}} \mathrm{diag}(e^{j\phi_{tx,0}}, e^{-j\phi_{tx,0}}, e^{-j\phi_{tx,0}}, e^{j\phi_{tx,0}}). \quad (28)$$

In (25), the number of nonlinear parameters to be estimated is reduced to 2, relaxing the computational demands of the estimation method.

1) Separable Nonlinear Least Squares: A natural solution to estimate the parameters is to minimize the squared error between $\mathbf{y}$ and $\mathbf{A}(\boldsymbol{\beta})\boldsymbol{\theta}_{\phi}$. Using this approach, the estimated parameters, denoted $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}_{\phi}$, can be obtained as:

$$\{\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}_{\phi}\} = \arg\min_{\boldsymbol{\beta}, \boldsymbol{\theta}_{\phi}} \|\mathbf{y} - \mathbf{A}(\boldsymbol{\beta})\boldsymbol{\theta}_{\phi}\|^2. \quad (29)$$

This approach is known as the Nonlinear Least Squares. For this particular signal model, it has been demonstrated that the minimization can be decoupled [47], [48]. Specifically, the minimizing argument can be obtained in two steps:

  a) Estimation of nonlinear parameters $\boldsymbol{\beta}$ as:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \|\mathbf{f}_{\boldsymbol{\beta}}\|^2. \quad (30)$$

  where $\mathbf{f}_{\boldsymbol{\beta}}$ is a function that computes the vector of residuals defined as:

$$\mathbf{f}_{\boldsymbol{\beta}} = \mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\beta})\mathbf{y}. \quad (31)$$

  with $\mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\beta}) = \mathbf{I}_{N_p} - \mathbf{P}_{\mathbf{A}}(\boldsymbol{\beta})$, while $\mathbf{I}_{N_p}$ is an $N_p \times N_p$ identity matrix, and $\mathbf{P}_{\mathbf{A}}(\boldsymbol{\beta})$ is the projection matrix onto the column space of $\mathbf{A}(\boldsymbol{\beta})$. The projection matrix $\mathbf{P}_{\mathbf{A}}(\boldsymbol{\beta})$ is defined as $\mathbf{P}_{\mathbf{A}}(\boldsymbol{\beta}) = \mathbf{A}(\boldsymbol{\beta})\mathbf{A}^{\dagger}(\hat{\boldsymbol{\beta}})$, where $\mathbf{A}^{\dagger}(\hat{\boldsymbol{\beta}})$ is the pseudoinverse of $\mathbf{A}(\boldsymbol{\beta})$ and is expressed as $\mathbf{A}^{\dagger}(\hat{\boldsymbol{\beta}}) = (\mathbf{A}^H(\boldsymbol{\beta})\mathbf{A}(\boldsymbol{\beta}))^{-1}\mathbf{A}^H(\boldsymbol{\beta})$,

  b) Estimation of the linear parameters $\boldsymbol{\theta}_{\phi}$ as

$$\hat{\boldsymbol{\theta}}_{\phi} = \mathbf{A}^{\dagger}(\hat{\boldsymbol{\beta}})\mathbf{y}. \quad (32)$$

2) Optimization algorithm: Considering that the most difficult operation of the estimation algorithm is the minimization problem in (30), and that the nonlinear parameters' values can vary in a large interval, a hybrid approach is employed to find the minimizing argument. Firstly, a grid search over the two-dimensional space related to $\boldsymbol{\beta}$ is performed. Secondly, using the value obtained as an initial guess, a local optimization algorithm is employed. For local optimization, we used the Levenberg–Marquardt (LM) iterative approach based on [49]. The LM algorithm is an optimization method that interpolates between the gradient descent and Gauss-Newton (GN) algorithms, being robust and converging almost as fast

as GN algorithm. Starting by using an initial guess $\boldsymbol{\beta_0} = \begin{bmatrix} \mathrm{Dz}_0 & \Delta f_0 \end{bmatrix}$ resulted from the grid search, the parameters update is obtained as:

$$\boldsymbol{\beta_{k+1}}^T = \boldsymbol{\beta_k}^T - [\mathbf{J}_{\boldsymbol{\beta}}^T \mathbf{J}_{\boldsymbol{\beta}} + \lambda_{\boldsymbol{\beta}} \mathbf{I}_2]^{-1} \mathbf{J}_{\boldsymbol{\beta}}^T \mathbf{f}_{\boldsymbol{\beta}} \big|_{\boldsymbol{\beta}=\boldsymbol{\beta_k}}. \quad (33)$$

where $\mathbf{J}_{\boldsymbol{\beta}}$ is the $N_p \times 2$ Jacobian matrix of the $\mathbf{f}_{\boldsymbol{\beta}}$ function, $\mathbf{I}_2$ is a 2×2 identity matrix, and $\lambda_{\boldsymbol{\beta}}$ is a damping parameter that controls the update step. The Jacobian matrix of the function $\mathbf{f}_{\boldsymbol{\beta}}$ can be expressed as follows:

$$\mathbf{J}_{\boldsymbol{\beta}} = \begin{bmatrix} \frac{\partial \mathbf{f}_{\boldsymbol{\beta}}}{\partial \beta_0} & \frac{\partial \mathbf{f}_{\boldsymbol{\beta}}}{\partial \beta_1} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\beta})}{\partial \beta_0} \mathbf{y} & \frac{\partial \mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\beta})}{\partial \beta_1} \mathbf{y} \end{bmatrix}. \quad (34)$$

where the derivative of the orthogonal projector is computed as in [50]:

$$\frac{\partial \mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\beta})}{\partial \beta_i} = -\mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\beta}) \frac{\partial \mathbf{A}(\boldsymbol{\beta})}{\partial \beta_i} \mathbf{A}^{\dagger}(\boldsymbol{\beta}) - \left(\mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\beta}) \frac{\partial \mathbf{A}(\boldsymbol{\beta})}{\partial \beta_i} \mathbf{A}^{\dagger}(\boldsymbol{\beta}) \right)^H. \quad (35)$$

and β_i is the i-th element of $\boldsymbol{\beta}$.

B. Compensation algorithm

The compensation algorithm aims to estimate the transmitted symbols from a block of N_b received samples. Considering that a block of data is generally longer than the training sequence ($N_b \gg N_p$), a phase tracking algorithm must be employed. In this objective, several L pilot symbols are periodically extracted from each of these data blocks using an $L \times N_b$ pilot extraction matrix denoted as $\mathbf{G}_p$. The pilots are typical symbols from whatever the modulation has been employed and can be expressed as $\mathbf{s}_p = \mathbf{G}_p \mathbf{s}$.

Once the parameters $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}_\phi$ have been estimated, the equation (21) can be rewritten as:

$$\mathbf{y} = \mathbf{A}(\hat{\boldsymbol{\beta}}, \phi) \mathbf{D}_\phi^{-1} \hat{\boldsymbol{\theta}}_\phi + \mathbf{b}. \quad (36)$$

Using the fact that $\mathbf{D}_\phi^{-1} = \mathbf{D}_\phi^*$, the impact of $\mathbf{D}_\phi^{-1}$ can be accumulated to the matrices $\boldsymbol{\Phi}(\phi_{tx})$ and $\boldsymbol{\Phi}(\phi_{rx})$ by using:

$$\mathbf{A}(\hat{\boldsymbol{\beta}}, \phi) \mathbf{D}_\phi^{-1} = \mathbf{A}(\hat{\boldsymbol{\beta}}, \boldsymbol{\varphi}) = \begin{bmatrix} \boldsymbol{\Phi}(\boldsymbol{\varphi_{rx}}) \mathbf{H}(\hat{\boldsymbol{\beta}}) \boldsymbol{\Phi}(\boldsymbol{\varphi_{tx}}) \mathbf{s} \\ \boldsymbol{\Phi}(\boldsymbol{\varphi_{rx}}) \mathbf{H}(\hat{\boldsymbol{\beta}}) \boldsymbol{\Phi}^*(\boldsymbol{\varphi_{tx}}) \mathbf{s}^* \\ \boldsymbol{\Phi}(\boldsymbol{\varphi_{rx}}) \mathbf{H}^*(\hat{\boldsymbol{\beta}}) \boldsymbol{\Phi}^*(\boldsymbol{\varphi_{tx}}) \mathbf{s}^* \\ \boldsymbol{\Phi}(\boldsymbol{\varphi_{rx}}) \mathbf{H}^*(\hat{\boldsymbol{\beta}}) \boldsymbol{\Phi}(\boldsymbol{\varphi_{tx}}) \mathbf{s} \end{bmatrix}^T. \quad (37)$$

where:

$$\boldsymbol{\Phi}(\boldsymbol{\varphi_{tx}}) = e^{-j\phi_{tx,0}} \boldsymbol{\Phi}(\boldsymbol{\phi_{tx}}), \quad (38a)$$

$$\boldsymbol{\Phi}(\boldsymbol{\varphi_{rx}}) = e^{-j\phi_{rx,0}} \boldsymbol{\Phi}(\boldsymbol{\phi_{rx}}). \quad (38b)$$

with $\boldsymbol{\varphi}_{tx} = \phi_{tx} - \phi_{tx,0}$, and $\boldsymbol{\varphi}_{rx} = \phi_{rx} - \phi_{rx,0}$.

Next, in order to detect the transmitted symbols using both the contribution of $\mathbf{s}$ and $\mathbf{s}^*$, we propose to linearly express the real and imaginary parts of $\mathbf{y}$ with respect to the real and imaginary parts of $\mathbf{s}$. Thus, the augmented model of the signal can be expressed as:

$$\tilde{\mathbf{y}} = \tilde{\boldsymbol{\Phi}}(\boldsymbol{\varphi_{rx}}) \tilde{\mathbf{M}} \tilde{\boldsymbol{\Phi}}(\boldsymbol{\varphi_{tx}}) \tilde{\mathbf{s}} + \tilde{\mathbf{b}}. \quad (39)$$

where

- $\tilde{\mathbf{M}}$ is an $2N_b \times 2N_b$ augmented matrix which is defined as:

$$\tilde{\mathbf{M}} = \begin{bmatrix} \Re e(\mathbf{M}_1) & -\Im m(\mathbf{M}_2) \\ \Im m(\mathbf{M}_1) & \Re e(\mathbf{M}_2) \end{bmatrix}. \quad (40)$$

with:

$$\mathbf{M}_1 = (\boldsymbol{\theta}_\phi^T (\mathbf{e}_0 + \mathbf{e}_1)) \mathbf{H}(\boldsymbol{\beta}) + (\boldsymbol{\theta}_\phi^T (\mathbf{e}_2 + \mathbf{e}_3)) \mathbf{H}^*(\boldsymbol{\beta}), \quad (41a)$$

$$\mathbf{M}_2 = (\boldsymbol{\theta}_\phi^T (\mathbf{e}_0 - \mathbf{e}_1)) \mathbf{H}(\boldsymbol{\beta}) + (\boldsymbol{\theta}_\phi^T (\mathbf{e}_3 - \mathbf{e}_2)) \mathbf{H}^*(\boldsymbol{\beta}). \quad (41b)$$

while $\mathbf{e}_k$ is the unit column vector that contains only an 1 value at the k-th row and 0s elsewhere,

- $\tilde{\boldsymbol{\Phi}}(\boldsymbol{\varphi_{tx}})$ and $\tilde{\boldsymbol{\Phi}}(\boldsymbol{\varphi_{rx}})$ are $2N_b \times 2N_b$ augmented matrices defined similarly to $\tilde{\boldsymbol{\Phi}}(\boldsymbol{\varphi})$ as follows:

$$\tilde{\boldsymbol{\Phi}}(\boldsymbol{\varphi}) = \begin{bmatrix} \Re e(\boldsymbol{\Phi}(\boldsymbol{\varphi})) & -\Im m(\boldsymbol{\Phi}(\boldsymbol{\varphi})) \\ \Im m(\boldsymbol{\Phi}(\boldsymbol{\varphi})) & \Re e(\boldsymbol{\Phi}(\boldsymbol{\varphi})) \end{bmatrix}. \quad (42)$$

1) Detection algorithm: The detection of the transmitted data is done by minimizing the squared error between the transmitted and received signals and can be obtained in two steps:

a) Estimation of the lasers' phases by using L pilot symbols as:

$$\{\hat{\boldsymbol{\varphi}}\} = \arg \min_{\boldsymbol{\varphi}} \|\mathbf{f}_{\boldsymbol{\varphi}}\|^2. \quad (43)$$

where $\boldsymbol{\varphi} = \begin{bmatrix} \boldsymbol{\varphi}_{tx}^T & \boldsymbol{\varphi}_{rx}^T \end{bmatrix}^T$, and $\mathbf{f}_{\boldsymbol{\varphi}}$ is a function that computes the vector of residuals, which by using the fact that $\tilde{\boldsymbol{\Phi}}^{-1}(\boldsymbol{\varphi}) = \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi})$, is defined as:

$$\mathbf{f}_{\boldsymbol{\varphi}} = \tilde{\mathbf{s}}_p - \tilde{\mathbf{G}}_p \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{tx}}) \tilde{\mathbf{M}}^{-1} \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{rx}}) \tilde{\mathbf{y}}. \quad (44)$$

with $\tilde{\mathbf{G}}_p = \mathbf{I}_2 \otimes \mathbf{G}_p$,

b) Estimation of the transmitted symbols by using a ZF equalizer as:

$$\hat{s}[n] = (\mathbf{e}_n + j\mathbf{e}_{n+N_b})^T \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{tx}}) \tilde{\mathbf{M}}^{-1} \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{rx}}) \tilde{\mathbf{y}}. \quad (45)$$
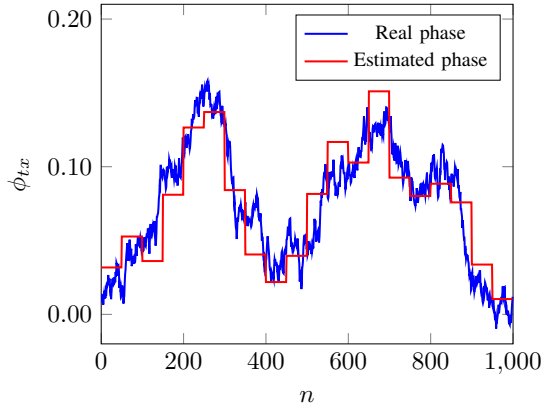
The constellation symbols can be finally estimated as follows:

$$\hat{s}_{\mathcal{M}}[n] = \arg \min_{s \in \mathcal{M}} \|s - \hat{s}[n]\|^2. \quad (46)$$

where $\hat{s}_{\mathcal{M}}[n]$ represents the n-th estimated constellation symbol.

2) Optimization algorithm: Considering the squared error function in (43), we start by setting up the initial guess $\boldsymbol{\varphi}_0 = \mathbf{0}_L$, where by $\mathbf{0}_L$ we denote a column vector of size L containing only 0s. Then, the parameters update is employed similarly to the one in equation (33). The $2L \times L$ Jacobian matrix of $\mathbf{f}_{\boldsymbol{\varphi}}$ can be expressed as:

$$\mathbf{J}_{\boldsymbol{\varphi}} = \begin{bmatrix} \frac{\partial \mathbf{f}_{\boldsymbol{\varphi}}}{\partial \varphi_{tx,0}} & \cdots & \frac{\partial \mathbf{f}_{\boldsymbol{\varphi}}}{\partial \varphi_{tx,L/2-1}} & \frac{\partial \mathbf{f}_{\boldsymbol{\varphi}}}{\partial \varphi_{rx,L/2}} & \cdots & \frac{\partial \mathbf{f}_{\boldsymbol{\varphi}}}{\partial \varphi_{rx,L-1}} \end{bmatrix}. \quad (47)$$

(a) Transmitter laser phase tracking
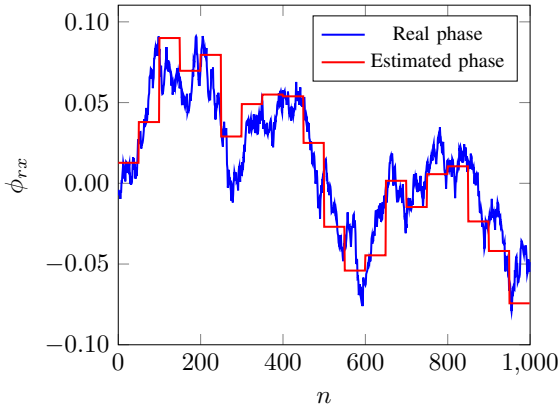
(b) Receiver laser phase tracking

Fig. 3: Lasers' phases joint estimation in a 16-QAM communication with infinite OSNR

where:

$$\frac{\partial \mathbf{f}_{\boldsymbol{\varphi}}}{\partial \varphi_{tx,i}} = -\tilde{\mathbf{G}}_p \frac{\partial \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{tx}})}{\partial \varphi_{tx,i}} \tilde{\mathbf{M}}^{-1} \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{rx}})\tilde{\mathbf{y}}, \qquad (48a)$$

$$\frac{\partial \mathbf{f}_{\boldsymbol{\varphi}}}{\partial \varphi_{rx,i}} = -\tilde{\mathbf{G}}_p \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{tx}}) \tilde{\mathbf{M}}^{-1} \frac{\partial \tilde{\boldsymbol{\Phi}}^T(\boldsymbol{\varphi_{rx}})}{\partial \varphi_{rx,i}}\tilde{\mathbf{y}}. \qquad (48b)$$

Considering that our tracking algorithm estimates φ and not the actual phases of the lasers ϕ, generally, a constant difference between the real and estimated phase will be observed. In order to emphasize the functionality of the proposed algorithm, we present in figure 3 a single realization of the joint estimation of the transmitter and receiver lasers' phases for a 100 kHz laser linewidth in a 16-QAM communication with infinite OSNR, by removing this constant difference.

IV. SIMULATION RESULTS AND DISCUSSIONS

In this section, the performances of the proposed algorithms are reported. The implementation is done based on (21), by using NumPy [51], and SciPy [52] Python libraries, while the LM algorithm is performed by calling a wrapper over Least Squares algorithms from MINPACK [53]. To assess the performance of our method, we used Monte Carlo simulations. Several metrics were employed to evaluate the performance,

including the MSE (Mean Squared Error) for estimation and the BER (Bit Error Rate) for compensation.

In the following simulations, we considered an $\mathcal{M}$-QAM coherent optical communication at 30 Gbaud symbol rate, with no oversampling employed, and transmitted over 1000 km of fiber on the wavelength of 1550 nm. The communication was impaired by the chromatic dispersion of the fiber which was assumed to have a 17 $ps/nm/km$ dispersion coefficient, 2 GHz CFO, 1 dB and 10° transmitter and receiver IQ imbalances, while multiple values between 100 kHz and 500 kHz were considered for the transmitter and receiver lasers' linewidth.

A. Estimation results

For the estimation of the quasi-static parameters, we used a number of N_p $\mathcal{M}$-QAM symbols. The nonlinear parameters search interval was chosen as follows:

- the parameter Dz was searched in an interval corresponding to a fiber length between 950 km and 1050 km as, generally, we have an approximate prior knowledge regarding the fiber length, and by limiting the search interval, we reduce the computational requirements,
- the parameter Δf was searched in an interval between ±3 GHz as we do not have any prior knowledge regarding the CFO, and over the lifetime of a typical tunable laser, the CFO can be as large as ±2.5 GHz [54].

To evaluate the theoretical performance of the proposed Nonlinear Least Squares estimator (NLSE), we compared its MSE to the Cramér Rao Lower Bound (CRLB). The CRLB places a lower bound on the variance of any unbiased estimator and its accuracy depends directly on the Probability Density Function of the data. The CRLB is obtained by inverting the Fisher information matrix and its computation is detailed in appendix A. The comparison was made with respect to the nonlinear parameters Dz and Δf, as the linear parameters estimation depends on the performance of the estimation of
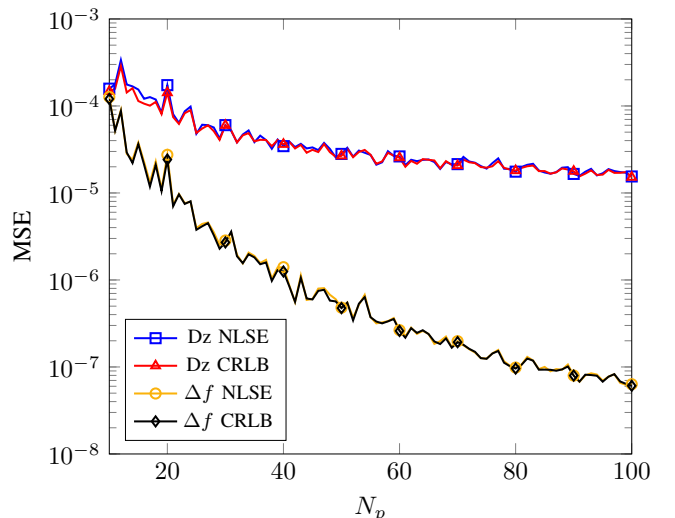

Fig. 4: MSE evolution of the NLSE and CRLB of nonlinear parameters for an ONSR of 20 dB with respect to N_p and by placing the noise after the receiver IQ imbalance
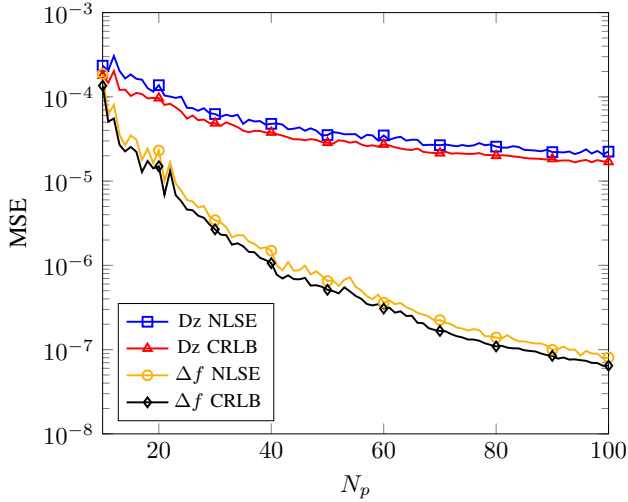
Fig. 5: MSE evolution of the NLSE and CRLB of nonlinear parameter for an ONSR of 20 dB with respect to N_p and by placing the noise before the receiver IQ imbalance

Fig. 6: MSE evolution of the NLSE of nonlinear parameters under the impact of PN, using $N_p = 100$ training symbols, with respect to the OSNR

the nonlinear parameters. In order to do that, we considered the ideal case where the lasers' phases are $0°$ ($\phi_{tx} = \phi_{rx} = 0$) over the entirety of the preamble symbols.

Regarding the noise placement in the communication chain, we considered two scenarios. Firstly, we considered the noise impact after the receiver IQ imbalance. In this scenario, the noise is circular. In figure 4, we presented the evolution of the MSE in this case, by varying N_p for an OSNR of 20 dB. The N_p variation is done with a step size of 1, while the curves' markers are displayed for a step size of 10. It can be observed that NLSE is optimal for both parameters as it asymptotically attains the CRLB. Also, it can be seen that Δf estimation has better performance in terms of MSE as compared to Dz estimation. Secondly, we considered the noise contribution before the receiver IQ imbalance. In this scenario, the noise loses the circularity property as the receiver IQ imbalance modifies its statistics. In figure 5, we presented the evolution of the MSE in this case, by varying N_p for an OSNR of 20 dB. Similarly to figure 4, the N_p variation is done with a step size of 1, and the curves' markers are displayed for a step size of 10. It can be observed that NLSE is not optimal as it does not attain the CRLB, but the MSE is relatively small for both parameters. Similarly to the first scenario, Δf estimation is better than the one of Dz, as the MSE is smaller. Starting from this point, in all the following simulations, the noise contribution is considered before the receiver IQ imbalance, as it corresponds to the worst-case estimation scenario.

In an experimental setup, the laser phase cannot be assumed null. Instead, it has a continuous slow variation compared to the signal phase. As a consequence, the estimation of the quasi-static parameters is impaired by the PN. We denoted the quasi-static nonlinear parameters estimator under the impact of PN as NLSE PN. Figure 6 presents the MSE evolution of the nonlinear parameters' estimation with respect to the OSNR for $N_p = 100$ preamble symbols. The results are compared to those corresponding to a scenario where the PN is null
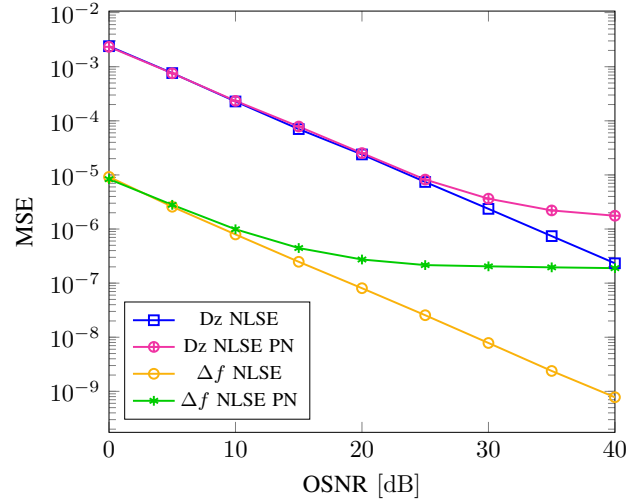
over the whole preamble. It can be seen that the PN has an important impact on the nonlinear parameters' estimation. For low values of OSNR, both estimators have similar results, but at high OSNRs, NLSE PN's performance is lower than the NLSE performance. This evolution is more visible for Δf.

B. Compensation performance

In this subsection, the performances are evaluated in the presence of the lasers' PN. For the estimation, we used a preamble of $N_p = 100$ symbols, and for the compensation, we used data blocks containing $N_b = 300$ $\mathcal{M}$-QAM symbols. From the N_b symbols, L of them are used as pilots to track the lasers' phases. A pre-FEC (Forward Error Correction) BER threshold (TH) of 3.8×10^{-3} was considered in order to obtain a post-FEC BER below 10^{-15} as indicated in Appendix I.9 of ITU-T G975.1 recommendation [55].
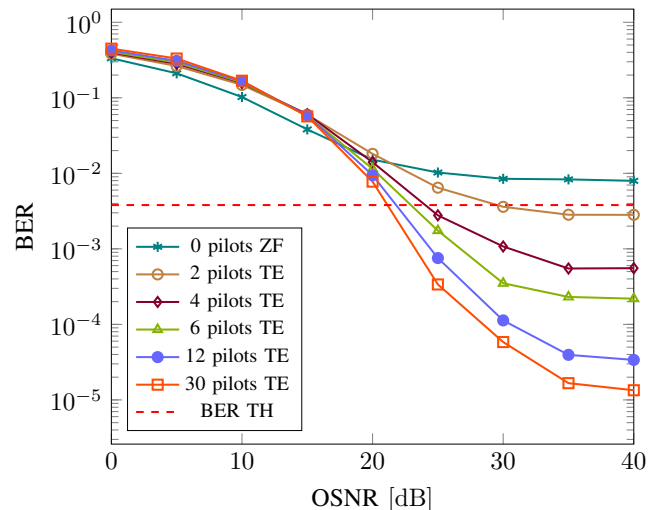

Fig. 7: BER evolution with respect to OSNR for a 16-QAM communication with $\delta f_{tx} = \delta f_{rx} = 100$ kHz and different L values for the pilots
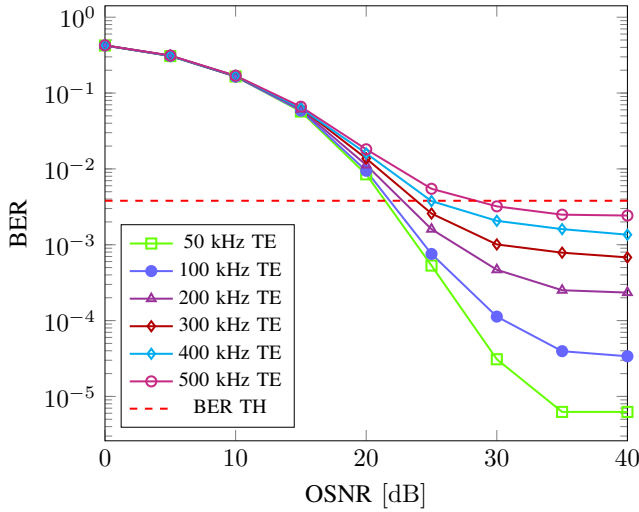
Fig. 8: BER evolution with respect to OSNR for a 16-QAM communication considering multiple $\delta f_{tx} = \delta f_{rx}$ values and by using $L = 12$ (4%) pilots
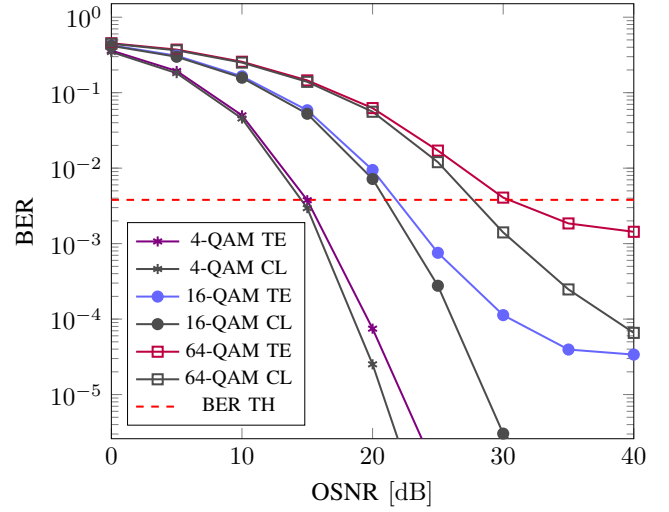

Fig. 9: BER evolution with respect to OSNR for different $\mathcal{M}$-QAM communications with $\delta f_{tx} = \delta f_{rx} = 100$ kHz and by using $L = 12$ (4%) pilots

In figure 7, the BER's evolution with respect to the OSNR for a 16-QAM coherent communication is shown. Multiple values of L pilots were employed, and a 100 kHz laser linewidth was used for both lasers ($\delta f_{tx} = \delta f_{rx} = 100$ kHz). It can be seen that the compensation performance is improved as the number of pilots is increased. At low ONSR values, the performances are similar, but as the OSNR increases, significant performance differences can be observed. We started by using a typical ZF equalizer that does not employ phase tracking. It can be seen that the system performance is strongly limited, and the BER has values above the BER TH even for high OSNR values. In order to improve the performance, we employed the proposed tracking algorithm before the ZF equalization, and we denoted this equalizer as TE (Tracking Equalizer). It can be seen that even by using 2 (0.67%) L pilots from the total N_b symbols, the BER is below BER TH starting from 30 dB OSNR. For the scenarios with and 4 (1.3%), 6 (2%), 12 (4%), and 30 (10%) pilots, above an OSNR value of 25 dB, the BER satisfies the targeted quality of service. Starting from this point, the $L = 12$ (4%) pilots scenario is used as it represents a good compromise between performance and overhead to data ratio.

In figure 8, the BER evolution is displayed for a 16-QAM coherent communication, with respect to OSNR and by considering different values for $\delta f_{tx} = \delta f_{rx}$. As expected, the performance of the TE decreases as the values of δf_{tx} and δf_{rx} increase. It can be observed that the BER curve is below BER TH starting from a 25 dB OSNR even for the case where $\delta f_{tx} = \delta f_{rx} = 400$ kHz. In the case where $\delta f_{tx} = \delta f_{rx} = 500$ kHz, the BER curve is below the BER TH for OSNRs greater than 30 dB.

In figure 9, BER's evolution is presented for 4-QAM, 16-QAM, 64-QAM communications with respect to OSNR. The performance is compared to a Clairvoyant (CL) equalizer [56] that assumes perfect knowledge of the quasi-static parameters on the receiver side, while the time-variant parameters are unknown and are estimated using pilots. The PN impact is

related to a laser linewidth of $\delta f_{tx} = \delta f_{rx} = 100$ kHz. It can be seen that the PN has a higher impact as the modulation order increases. At low OSNRs, the TE has similar values to the CL, but as the OSNRs increases, the performance of TE is more severely degraded by the PN. For a 4-QAM modulation, the BER values are below BER TH for OSNRs higher than 15 dB and a 0.5 dB OSNR penalty with respect to the CL is introduced. Regarding the 16-QAM communication, the BER values are below the BER TH for OSNR values larger than 21.8 dB, and a 0.8 dB OSNR penalty with respect to the CL is introduced at this point. It should be mentioned that the BER values are below the BER TH even for the 64-QAM communication for ONSRs higher than 30.4 dB. At this point, a 2.7 dB OSNR penalty with respect to the CL is introduced.

C. Complexity analysis

In this subsection, we performed an analysis of the computational complexity requirements of our method. The analysis takes into consideration the number of operations and the computational time required by the proposed method.

The estimation algorithm complexity analysis can be divided into two parts corresponding to the nonlinear parameters' estimation and linear parameters' estimation. The nonlinear parameters estimation is the most computationally demanding task since it employs a grid search over a two-dimensional space and an LM optimization, as opposed to the estimation of the linear parameter, which only computes the pseudoinverse of a matrix. In table I, the approximative number of operations required for each step is shown. As it can be seen, the computational complexity of the estimation algorithm approaches $\mathcal{O}(N_p^2)$ for $N_p \to \infty$, with the Jacobian matrix computation being the most demanding step. The convergence of the algorithm depends on the initial guess provided by the grid search. Assuming an initial guess close to the global minimum of the function, the LM optimization normally converges to the minimum of the cost function. The convergence of the algorithm was approximated as in [57].

This article has been accepted for publication in IEEE Systems Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSYST.2021.3111777

ACCEPTED TO THE IEEE SYSTEMS JOURNAL

9

Computation	Number of operations
Grid search	$48N_p^2 + 2N_p \log N_p + 70N_p$
Residual function	$80N_p^2 + 2N_p \log N_p + 36N_p$
Jacobian matrix	$178N_p^2 + 6N_p \log N_p + 35N_p$
Pseudoinverse matrix	$48N_p^2 + 4N_p \log N_p + 50N_p$

TABLE I: Approximative number of operations needed for the estimation algorithm

Initially, the LM algorithm converges linearly, then the order of convergence increases to values between 1 and 2, and at the solution, the convergence returns to a linear evolution. Similarly, the compensation algorithm complexity analysis can be divided into two parts. The first part corresponds to the phase tracking, and the second one to the ZF equalization and constellation enforcement. The phase tracking algorithm employs LM optimization. The ZF equalizer is computed using a matrix inversion, while the constellation enforcement is based on the $\arg\min$ operation. In table II, the approximative number of operations required for each step is shown. As it can be seen, the computational complexity of the compensation algorithm approaches $\mathcal{O}(N_b^3)$ for $N_p \to \infty$, with the ZF equalization being the most demanding step. As the cost function is a convex function, the algorithm normally converges independent of the initial guess. Similar to the estimation algorithm, the LM algorithm starts converging linearly, then the order of convergence increases to values between 1 and 2, and at the solution, the convergence returns to a linear evolution.

For the simulations, it was used a virtual machine with an Intel(R) Xeon(R) Platinum 8171M CPU. The CPU has 2 cores, 4 threads and operates at a frequency of 2.60 GHz. The machine has 16 GB RAM and a 32 GB SSD. Regarding the estimation, for a number of $N_p = 100$ training symbols, and an OSNR of 20 dB, 10 Dz and 15 Δf uniformly distributed values are needed to find a reliable initial guess using the grid search. In this context, 150 evaluations of a cost function, an average of 6 evaluations of the residual function $\mathbf{f}_\beta$, and 5 evaluations of the Jacobian $\mathbf{J}_\beta$ were computed in order for the algorithm to converge. In this scenario, the average running time needed for the estimation is 0.095 s. It is worth mentioning that the grid search has the highest impact on computational complexity. Regarding the compensation, for a number of $N_b = 300$ training symbols, an OSNR of 20 dB, and $L = 12$ pilots, an average of 7 evaluations of the residual function $\mathbf{f}_\varphi$, and 6 evaluations of the Jacobian $\mathbf{J}_\varphi$ were computed in order for the algorithm to converge. In this scenario, the average running time needed for the compensation is 0.091 s. The total average running time for a full communication chain simulation using the parameters mentioned above is 0.224 s.

Computation	Number of operations
Residual function	$4N_b^2 + 12N_b + 2L$
Jacobian matrix	$16N_b^2 L + 48N_b L$
ZF Equalizer	$12N_b^3 + 37N_b^2 + 13N_b$

TABLE II: Approximative number of operations needed for the compensation algorithm

V. CONCLUSION

In this paper, we have proposed an estimation and compensation technique capable of mitigating several linear effects that impair coherent optical systems. The proposed algorithm jointly estimates and compensates for CD, CFO, transmitter and receiver IQ imbalance and PN. Results from numerical simulations have demonstrated the efficiency of our method for $\mathcal{M}$-QAM modulation formats. It was shown that the proposed algorithm could estimate the quasi-static nonlinear system parameters for a wide range of values. In an ideal scenario with no PN, the estimation performance is close to the theoretical performance provided by the CRLB. In a more realistic scenario, although limited by the PN presence, the estimation still offers good MSE performances. A phase tracking algorithm based on pilots and LM optimization was employed. It was shown, in particular, that it can jointly track both transmitter and receiver lasers' phases. The compensation algorithm shows good performance in terms of BER in the presence of multiple linear impairments. Moreover, it was shown that the computational time is limited for both algorithms.

In future work, we will propose to extend the algorithm for coherent polarization division-multiplexed systems as the polarization mode dispersion parameters estimation may be inserted in the optimization algorithm of the quasi-static parameters. Moreover, we consider studying the possibility of adapting the proposed algorithm for OFDM communications.

APPENDIX

In this appendix, we show how to compute the CRLB to theoretically validate the estimator performance, as for any unbiased estimator, the MSE is lower bounded by the CRLB [48].

Starting from the equation (25), it can be seen that when all the parameters from the vector $\boldsymbol{\theta}_\phi$ are unknown, the signal model contains parameters indetermination. For this reason, firstly, we consider $\phi_{tx} = \phi_{rx} = 0$, and in this case $\boldsymbol{\theta}_\phi \to \boldsymbol{\theta}$. Secondly, without lack of generality and leaving the noise statistics unchanged, we assume that $\mu_{rx} = 1$. Furthermore, for practical considerations, we also exclude the unrealistic case $|\nu_{rx}| \geq 1$ from the estimation problem.

Under these assumptions, the received signal can be written as:

$$\mathbf{y} = \mathbf{s}(\boldsymbol{\Omega}) + \mathbf{b}. \tag{49}$$

where $\boldsymbol{\Omega}$ is a real-valued vector containing the parameters to estimate, $\mathbf{s}(\boldsymbol{\Omega})$ is a vector containing the deterministic part of the signal defined as:

$$\mathbf{s}(\boldsymbol{\Omega}) = \mathbf{F}(\boldsymbol{\beta})\boldsymbol{\theta}_{tx} + \nu_{rx}\mathbf{F}^*(\boldsymbol{\beta})\boldsymbol{\theta}_{tx}^*. \tag{50}$$

with

$$\mathbf{F}(\boldsymbol{\beta}) = \mathbf{H}(\boldsymbol{\beta})\mathbf{S}, \tag{51}$$

$$\boldsymbol{\theta}_{tx} = \begin{bmatrix} \mu_{tx} \\ \nu_{tx} \end{bmatrix}, \tag{52}$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{s} & \mathbf{s}^* \end{bmatrix}. \tag{53}$$

and $\mathbf{b}$ is a vector containing the noise samples.

In order to compute the CRLB, we express the augmented real-valued signal model as follows:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{s}}(\boldsymbol{\Omega}) + \tilde{\mathbf{b}}. \tag{54}$$

where $\tilde{\mathbf{s}}(\boldsymbol{\Omega})$ is given by:

$$\tilde{\mathbf{s}}(\boldsymbol{\Omega}) = \begin{bmatrix} \Re e((1 + \nu_{rx})\mathbf{F}(\boldsymbol{\beta})\boldsymbol{\theta}_{tx}) \\ \Im m((1 - \nu_{rx})\mathbf{F}(\boldsymbol{\beta})\boldsymbol{\theta}_{tx}) \end{bmatrix}. \tag{55}$$

In this context, the number of real-valued parameters to be estimated is equal to 9. These parameters are given by:

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\beta} & \boldsymbol{\Omega}_{rx} & \boldsymbol{\Omega}_{tx} & \sigma^2 \end{bmatrix}^T. \tag{56}$$

where

- $\boldsymbol{\Omega}_r = \begin{bmatrix} \Re e(\nu_{rx}) & \Im m(\nu_{rx}) \end{bmatrix}$ is a row vector containing the real and imaginary parts of the receiver IQ imbalance parameter,
- $\boldsymbol{\Omega}_t = \begin{bmatrix} \Re e(\boldsymbol{\theta}_{tx}^T) & \Im m(\boldsymbol{\theta}_{tx}^T) \end{bmatrix}$ is a row vector containing the real and imaginary parts of the transmitter IQ imbalance parameters,
- σ^2 is the noise variance.

Let us denote by Ω_k the k-th element of the vector $\boldsymbol{\Omega}$. The CRLBs of Ω_k is given by the k-th diagonal element of the inverse of the Fisher Information Matrix i.e.:

$$\mathrm{CRLB}[\Omega_k] = [\mathbf{I}^{-1}(\boldsymbol{\Omega})]_{kk}. \tag{57}$$

where $\mathbf{I}(\boldsymbol{\Omega})$ is the Fisher Information Matrix, and $[.]_{kl}$ corresponds to the (k, l)-th element of a matrix. As the augmented received vector is distributed as $\tilde{\mathbf{y}} \sim \mathcal{N}(\tilde{\mathbf{s}}(\boldsymbol{\Omega}), \mathbf{C}(\boldsymbol{\Omega}))$, where $\mathbf{C}(\boldsymbol{\Omega})$ is the covariance matrix, the (k, l)-th element of the Fisher Information Matrix is given by [48]:

$$[\mathbf{I}(\boldsymbol{\Omega})]_{kl} = \left[\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_k}\right]^T \mathbf{C}^{-1}(\boldsymbol{\Omega}) \left[\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_l}\right]$$
$$+ \frac{1}{2}\mathbf{tr}\left[\mathbf{C}^{-1}(\boldsymbol{\Omega})\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_k}\mathbf{C}^{-1}(\boldsymbol{\Omega})\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_l}\right]. \tag{58}$$

Regarding the partial derivatives of $\tilde{\mathbf{s}}(\boldsymbol{\Omega})$, $\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial \sigma^2} = \mathbf{0}$, and:

$$\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\beta}]_k} = \begin{bmatrix} \Re e((1 + \nu_{rx}^*)\mathbf{G}_k\boldsymbol{\theta}_{tx}) \\ \Im m((1 - \nu_{rx}^*)\mathbf{G}_k\boldsymbol{\theta}_{tx}) \end{bmatrix},$$
$$\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_{rx}]_k} = \begin{bmatrix} \Re e(\mathbf{F}(\boldsymbol{\beta})\boldsymbol{\theta}_{tx}) & \Im m(\mathbf{F}(\boldsymbol{\beta})\boldsymbol{\theta}_{tx}) \\ -\Im m(\mathbf{F}(\boldsymbol{\beta})\boldsymbol{\theta}_{tx}) & \Re e(\mathbf{F}(\boldsymbol{\beta})\boldsymbol{\theta}_{tx}) \end{bmatrix}\mathbf{e}_k,$$
$$\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_{tx}]_k} = \begin{bmatrix} \Re e((1 + \nu_{rx}^*)\mathbf{F}(\boldsymbol{\beta})) \\ \Im m((1 - \nu_{rx}^*)\mathbf{F}(\boldsymbol{\beta})) \\ -\Im m((1 + \nu_{rx}^*)\mathbf{F}(\boldsymbol{\beta})) \\ \Re e((1 - \nu_{rx}^*)\mathbf{F}(\boldsymbol{\beta})) \end{bmatrix}\mathbf{e}_k.$$

where:

$$\mathbf{G}_k = \frac{\partial \mathbf{F}(\boldsymbol{\beta})}{\partial [\boldsymbol{\beta}]_k} = \frac{\partial \mathbf{H}(\boldsymbol{\beta})}{\partial [\boldsymbol{\beta}]_k}\mathbf{S}. \tag{59}$$

The covariance matrix derivatives computation depends on the noise contribution placement. Firstly, we consider the scenario where the noise is placed after the receiver IQ imbalance. In this case, as the noise is circular, $\mathbf{C}(\boldsymbol{\Omega}) = \frac{\sigma^2}{2}\mathbf{I}_{2N_p}$, and the

derivatives $\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_k}$ are non-zero only for the noise variance σ^2. This non-zero derivative is given by:

$$\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial \sigma^2} = \frac{1}{2}\mathbf{I}_{2N_p}.$$

Secondly, we consider the scenario where the noise is placed before the receiver IQ imbalance. In this case, the noise is non-circular, and $\tilde{\mathbf{b}}$ is defined as:

$$\tilde{\mathbf{b}} = (\tilde{\mathbf{L}}(\nu_{rx}) \otimes \mathbf{I}_{N_p})\tilde{\mathbf{w}}. \tag{60}$$

with:

$$\tilde{\mathbf{L}}(\nu_{rx}) = \mathbf{I}_2 + \begin{bmatrix} \Re e(\nu_{rx}) & \Im m(\nu_{rx}) \\ \Im m(\nu_{rx}) & -\Re e(\nu_{rx}) \end{bmatrix}. \tag{61}$$

and $\mathbf{w}$ being the circular noise before the receiver IQ imbalance. The covariance matrix is expressed as $\mathbf{C}(\boldsymbol{\Omega}) = \frac{\sigma^2}{2}\tilde{\mathbf{L}}(\nu_{rx})\tilde{\mathbf{L}}^T(\nu_{rx}) \otimes \mathbf{I}_{N_p}$, and the derivatives $\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_k}$ are non-zero only for the receiver's IQ parameter $\boldsymbol{\Omega}_{rx}$, and for the noise variance σ^2. These non-zero derivatives are given by:

$$\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_{rx}]_0} = \sigma^2 \begin{bmatrix} \Re e(\nu_{rx}) + 1 & 0 \\ 0 & \Re e(\nu_{rx}) - 1 \end{bmatrix} \otimes \mathbf{I}_{N_p},$$
$$\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_{rx}]_1} = \sigma^2 \begin{bmatrix} \Im m(\nu_{rx}) & 1 \\ 1 & \Im m(\nu_{rx}) \end{bmatrix} \otimes \mathbf{I}_{N_p},$$
$$\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial \sigma^2} = \frac{1}{2}\tilde{\mathbf{L}}(\nu_{rx})\tilde{\mathbf{L}}^T(\nu_{rx}) \otimes \mathbf{I}_{N_p}.$$

REFERENCES

[1] Cisco (2020, Mar), Cisco Annual Internet Report (2018–2023) White Paper [Online]. Avalaible: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf.

[2] T.-H. Nguyen, P. Scalart, M. Gay, L. Bramerie, O. Sentieys, J.-C. Simon, C. Peucheret, and M. Joindot, "Blind transmitter IQ imbalance compensation in $\mathcal{M}$-QAM optical coherent systems," *Journal of Optical Communications and Networking*, vol. 9, no. 9, pp. D42–D50, Sept 2017.

[3] C. R. S. Fludger and T. Kupfer, "Transmitter impairment mitigation and monitoring for high baud-rate, high order modulation systems," in *ECOC 2016; 42nd European Conference on Optical Communication*, Dusseldorf, Germany, 2016, pp. 1–3.

[4] T. Xu, G. Jacobsen, S. Popov, J. Li, E. Vanin, K. Wang, A. T. Friberg, and Y. Zhang, "Chromatic dispersion compensation in coherent transmission system using digital filters," *Optics Express*, vol. 18, no. 15, pp. 16 243–16 257, July 2010.

[5] A. Meiyappan, P.-Y. Kam, and H. Kim, "On decision aided carrier phase and frequency offset estimation in coherent optical receivers," *IEEE Journal of Lightwave Technology*, vol. 31, no. 13, pp. 2055–2069, July 2013.

[6] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning," *IEEE Photonics Technology Letters*, vol. 22, no. 9, pp. 631–633, May 2010.

[7] A. Ellis, M. McCarthy, M. Al Khateeb, M. Sorokina, and N. Doran, "Performance limits in optical communications due to fiber nonlinearity," *Advances in Optics and Photonics*, vol. 9, no. 3, pp. 429–503, July 2017.

[8] E. M. Ip and J. M. Kahn, "Fiber impairment compensation using coherent detection and digital signal processing," *IEEE Journal of Lightwave Technology*, vol. 28, no. 4, pp. 502–519, Feb 2010.

[9] K. Kikuchi, "Fundamentals of coherent optical fiber communications," *IEEE Journal of Lightwave Technology*, vol. 34, no. 1, pp. 157–179, Jan 2015.

[10] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," *IEEE Journal of Lightwave Technology*, vol. 35, no. 5, pp. 1125–1141, Mar 2017.

[11] I. Fatadin, S. J. Savory, and D. Ives, "Compensation of quadrature imbalance in an optical QPSK coherent receiver," *IEEE Photonics Technology Letters*, vol. 20, no. 20, pp. 1733–1735, Oct 2008.

[12] S. H. Chang, H. S. Chung, and K. Kim, "Impact of quadrature imbalance in optical coherent QPSK receiver," *IEEE Photonics Technology Letters*, vol. 21, no. 11, pp. 709–711, June 2009.

[13] M. S. Faruk and K. Kikuchi, "Compensation for in-phase/quadrature imbalance in coherent-receiver front end for optical quadrature amplitude modulation," *IEEE Photonics Journal*, vol. 5, no. 2, pp. 7 800 110–7 800 110, Apr 2013.

[14] E. P. da Silva and D. Zibar, "Widely linear equalization for IQ imbalance and skew compensation in optical coherent receivers," *IEEE Journal of Lightwave Technology*, vol. 34, no. 15, pp. 3577–3586, Aug 2016.

[15] M. S. Faruk and K. Kikuchi, "Front-end IQ-error compensation in coherent optical receivers," in *2012 17th Opto-Electronics and Communications Conference*, Busan, South Korea, 2012, pp. 156–157.

[16] Q. Xiang, Y. Yang, Q. Zhang, J. Cao, and Y. Yao, "Compensation of IQ imbalance using Kalman filter in coherent optical systems," in *2018 23rd Opto-Electronics and Communications Conference (OECC)*, Jeju, South Korea, 2018, pp. 1–2.

[17] M. K. Lagha, R. Gerzaguet, L. Bramerie, M. Gay, M.-L. Chares, C. Peucheret, and P. Scalart, "Blind joint polarization demultiplexing and IQ imbalance compensation for $\mathcal{M}$-QAM coherent optical communications," *IEEE Journal of Lightwave Technology*, vol. 38, no. 16, pp. 4213–4220, Aug 2020.

[18] Q. Zhang, Y. Yang, C. Guo, X. Zhou, Y. Yao, A. P. T. Lau, and C. Lu, "Algorithms for blind separation and estimation of transmitter and receiver IQ imbalances," *IEEE Journal of Lightwave Technology*, vol. 37, no. 10, pp. 2201–2208, May 2019.

[19] X. Dai, X. Li, M. Luo, and S. Yu, "Numerical simulation and experimental demonstration of accurate machine learning aided IQ time-skew and power-imbalance identification for coherent transmitters," *Optics express*, vol. 27, no. 26, pp. 38 367–38 381, Dec 2019.

[20] N. Stojanovic and X. Changsong, "An efficient method for skew estimation and compensation in coherent receivers," *IEEE Photonics Technology Letters*, vol. 28, no. 4, pp. 489–492, Feb 2016.

[21] J. C. M. Diniz, F. D. Ros, R. T. Jones, and D. Zibar, "Time skew estimator for dual-polarization QAM transmitters," in *2017 European Conference on Optical Communication (ECOC)*, Gothenburg, Sweden, 2017, pp. 1–3.

[22] X. Dai, M. Luo, and X. Li, "Machine learning aided in-phase/quadrature skew and imbalance calibration for coherent optical transmitters," in *2019 Optical Fiber Communications Conference and Exhibition (OFC)*, San Diego, CA, USA, 2019, pp. 1–3.

[23] G. P. Agrawal, *Fiber-optic Communication Systems*, 4th ed. John Wiley & Sons, 2010.

[24] A. Eghbali, H. Johansson, O. Gustafsson, and S. J. Savory, "Optimal least-squares FIR digital filters for compensation of chromatic dispersion in digital coherent optical receivers," *IEEE Journal of Lightwave Technology*, vol. 32, no. 8, pp. 1449–1456, Apr 2014.

[25] G. Goldfarb and G. Li, "Chromatic dispersion compensation using digital IIR filtering with coherent detection," *IEEE Photonics Technology Letters*, vol. 19, no. 13, pp. 969–971, July 2007.

[26] S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, "Electronic compensation of chromatic dispersion using a digital coherent receiver," *Opt. Express*, vol. 15, no. 5, pp. 2120–2126, Mar 2007. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-15-5-2120

[27] I. Fatadin and S. J. Savory, "Compensation of frequency offset for 16-QAM optical coherent systems using QPSK partitioning," *IEEE Photonics Technology Letters*, vol. 23, no. 17, pp. 1246–1248, Sept 2011.

[28] Q. Yan, L. Liu, and X. Hong, "Blind carrier frequency offset estimation in coherent optical communication systems with probabilistically shaped $\mathcal{M}$-QAM," *IEEE Journal of Lightwave Technology*, vol. 37, no. 23, pp. 5856–5866, Dec 2019.

[29] M. Li and L.-K. Chen, "Blind carrier frequency offset estimation based on eighth-order statistics for coherent optical QAM systems," *IEEE Photonics Technology Letters*, vol. 23, no. 21, pp. 1612–1614, Nov 2011.

[30] D. Huang, T.-H. Cheng, and C. Yu, "Accurate two-stage frequency offset estimation for coherent optical systems," *IEEE Photonics Technology Letters*, vol. 25, no. 2, pp. 179–182, Jan 2013.

[31] W. Jiang, Y. Yang, Q. Zhang, Y. Sun, K. Zhong, X. Zhou, and Y. Yao, "Application of Kalman filter in frequency offset estimation for coherent optical quadrature phase-shift keying communication system," *Optical Engineering*, vol. 55, no. 9, p. 096102, Sept 2016.

[32] A. Meiyappan, P.-Y. Kam, and H. Kim, "A complex-weighted, decision-aided, maximum-likelihood carrier phase and frequency-offset estimation algorithm for coherent optical detection," *Opt. Express*, vol. 20, no. 18, pp. 20 102–20 114, Aug 2012. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-18-20102

[33] T. Pfau, S. Hoffmann, and R. Noé, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *IEEE Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, Apr 2009.

[34] G. Colavolpe, T. Foggi, E. Forestieri, and M. Secondini, "Impact of phase noise and compensation techniques in coherent optical systems," *IEEE Journal of Lightwave Technology*, vol. 29, no. 18, pp. 2790–2800, Sept 2011.

[35] J. H. Ke, K. P. Zhong, Y. Gao, J. C. Cartledge, A. S. Karar, and M. A. Rezania, "Linewidth-tolerant and low-complexity two-stage carrier phase estimation for dual-polarization 16-QAM coherent optical fiber communications," *IEEE Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3987–3992, Dec 2012.

[36] A. Kakkar, J. R. Navarro, R. Schatz, H. Louchet, X. Pang, O. Ozolins, G. Jacobsen, and S. Popov, "Comprehensive study of equalization-enhanced phase noise in coherent optical systems," *Journal of Lightwave Technology*, vol. 33, no. 23, pp. 4834–4841, Dec 2015.

[37] W. Shieh and K.-P. Ho, "Equalization-enhanced phase noise for coherent-detection systems using electronic digital signal processing," *Opt. Express*, vol. 16, no. 20, pp. 15 718–15 727, Sept 2008. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-16-20-15718

[38] D. Huang, T.-H. Cheng, and C. Yu, "Decision-aided carrier phase estimation with selective averaging for low-cost optical coherent communication," in *2013 9th International Conference on Information, Communications & Signal Processing*. Tainan, Taiwan: IEEE, 2013, pp. 1–3.

[39] D. Tandur and M. Moonen, "Joint adaptive compensation of transmitter and receiver IQ imbalance under carrier frequency offset in OFDM-based systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5246–5252, Nov 2007.

[40] A. Tarighat, R. Bagheri, and A. H. Sayed, "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3257–3268, Aug 2005.

[41] A. Mehrabian and A. Zaimbashi, "Spectrum sensing in SIMO cognitive radios under primary user transmitter IQ imbalance," *IEEE Systems Journal*, vol. 13, no. 2, pp. 1210–1218, June 2018.

[42] F. Jalali and A. Zaimbashi, "Cognitive radio spectrum sensing under joint TX/RX I/Q imbalance and uncalibrated receiver," *IEEE Systems Journal*, vol. 14, no. 1, pp. 105–112, Mar 2019.

[43] S. J. Savory, "Digital filters for coherent optical receivers," *Optics Express*, vol. 16, no. 2, pp. 804–817, Jan 2008.

[44] F. Kalbat, A. Al-Dweik, B. Sharif, and G. K. Karagiannidis, "Performance analysis of precoded wireless OFDM with carrier frequency offset," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2237–2248, June 2019.

[45] S. J. Savory, "Digital Coherent Optical Receivers: Algorithms and Subsystems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1164–1179, Sept 2010.

[46] R. Borkowski, D. Zibar, and I. T. Monroy, "Anatomy of a digital coherent receiver," *IEICE Transactions on Communications*, vol. 97, no. 8, pp. 1528–1536, Aug 2014.

[47] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, no. 2, pp. 413–432, Apr 1973.

[48] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimtion Theory*. Prentice Hall PTR, 1993.

[49] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical Analysis*. Springer, 1978, pp. 105–116.

[50] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 11, pp. 2436–2449, Nov 1991.

[51] C. R. Harris *et al.*, "Array programming with NumPy," *Nature*, vol. 585, p. 357–362, Sept 2020.

[52] P. Virtanen *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, Feb 2020.

[53] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "User guide for MINPACK-1," CM-P00068642, Geneva, Switzerland, Tech. Rep., Aug 1980.

[54] O. I. Forum, *Integrable Tunable Laser Assembly Multi Source Agreement, OIF-ITLA-MSA-01.3*, July 2015.

[55] ITU-T G975.1 Recommandation, "Forward error correction for high bit-rate DWDM submarine systems," Feb 2004.

[56] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall PTR, 1998.

[57] J. Stoer and R. Bulirsch, *Introduction to numerical analysis*. Springer Science & Business Media, 2013.