Universal Virtual Lab: A Fast and Accurate Simulation Tool for Wideband Nonlinear DWDM Systems

David Dahan, *Senior Member, IEEE*, Michael Zarubinsky, Yunhua Liang, Ori Golani, and Mark Shtaif, *Fellow, IEEE, Fellow, OSA*

***Abstract*—We introduce the concept of the universal virtual lab, an extension to the virtual lab platform of [Golani *et al.* 2016], enabling a fast and accurate simulation of wideband nonlinear DWDM systems. The universal virtual lab is compliant with non-ideal transmitter and receiver architectures, distributed optical filters in the optical link, inter-channel stimulated Raman scattering, and it provides accurate performance predictions even when adaptive equalization methods are applied. In comparison with the conventional full-bandwidth split step Fourier transform method, we show with different test scenarios that the universal virtual lab provides accuracy errors below 0.1 dBQ and 0.09 bit/4D-symb in Q-factor and GMI assessments respectively, with runtime speedup factors exceeding 1000. We also report performance assessments in an ultra-wideband (11 THz) C+L system and discuss equalization gain under different compensation scenarios. The estimated speedup factor with respect to the full-bandwidth split step Fourier transform method is assessed to be greater than 35,000.**

***Index Terms*—Fiber nonlinear optics, nonlinear interference, nonlinearity compensation, optical fiber communication, stimulated Raman scattering, time-varying inter symbol interference, ultra-wideband systems.**

I. Introduction

THE exponentially increasing data traffic in optical fiber networks has pushed towards the deployment of new technologies such as wideband dense wavelength division multiplexing (DWDM) systems with C+L and S+C+L band operations over more than 10 THz bandwidth [1]–[3] and spectrally efficient optical coherent modulation formats [4], [5]. Optical network throughputs are mainly bounded by the nonlinear Shannon capacity [6], with optical channels being distorted by Kerr induced nonlinearities such as self-phase modulation (SPM), cross-phase phase modulation (XPM) and four-wave mixing (FWM) [7][1]. For wideband systems, inter-channel stimulated Raman scattering (ISRS) leads to additional penalties as some portions of the channel optical power are transferred from higher to lower optical frequencies via nonlinear inelastic scattering of the signal photons by the fiber silica medium [8]. Accurate and reliable simulations of nonlinear DWDM systems with a large number of interfering channels (IC) are critical for designing and predicting the performance of fiber-communications systems [9]–[11], for the development of nonlinearity tolerant modulation formats and for devising effective nonlinearity equalization techniques [12]–[14], as well as for optical networking related applications such as optical channel monitoring [15], fast and accurate network planning tools [16] and means against physical layer related security threats [17]. Simulation of large channel-count systems are extremely demanding in terms of computational resources and often require unacceptably long computation time when relying on the conventional split-step Fourier-transform method (SSFTM) [18]. A GPU based implementation of the SSFTM has been reported in [19] but even then, the runtime was only decreased by a factor of 40 for a simulation bandwidth of 4 THz. Since the complexity grows at least as a fourth power of the DWDM bandwidth [18], SSFTM becomes impractical for ultra-wideband systems operating beyond the C-band and alternative solutions must be proposed while ensuring a careful trade-off between accuracy and computational runtime.

Throughout the last decade, several models have been developed to address nonlinear optical fiber induced distortions as nonlinear interference noise (NLIN) within the framework of a first order perturbation solution of the Manakov equation. They include the Gaussian noise (GN) model [20], the enhanced Gaussian noise model (EGN) [21], both developed in the frequency domain and the time-varying intersymbol interference (ISI) model of the XPM induced NLIN [22], [23]. Some of these models have also been extended to include the ISRS impact [11], [24]. These models lead to semi-analytical solutions of the NLIN and in some contexts, they offer an acceptable alternative to SSFTM simulations. While the EGN model provides a very

Manuscript received October 13, 2021; revised December 12, 2021 and January 4, 2022; accepted January 5, 2022. Date of publication January 10, 2022; date of current version April 18, 2022. *(Corresponding author: David Dahan.)*

David Dahan is with the Toga Networks, A Huawei Company, Hod Hasharon 45240, Israel, and also with the Department of Electrical Engineering, Holon Institute of Technology, Holon 58102, Israel (e-mail: david.dahan@huawei.com).

Michael Zarubinsky and Yunhua Liang are with the Toga Networks, A Huawei Company, Hod Hasharon 45240, Israel (e-mail: michael.zarubinsky@huawei.com; liangyunhua1@huawei.com).

Ori Golani and Mark Shtaif are with the School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel (e-mail: ori.golani@gmail.com; shtaif@tauex.tau.ac.il).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2022.3141447.

Digital Object Identifier 10.1109/JLT.2022.3141447

[1]These three Kerr induced effects are also referred in the literature as single-channel interference (SCI), cross-channel interference (XCI) and multi-channel interference (MCI), respectively.

good estimate of the NLIN variance, it does not account for the correlation properties of the NLIN and for this reason, it is primarily used for assessing the performance of systems in which adaptive equalization is absent [25]–[27]. The time varying ISI model, on the other hand, accounts for temporal correlation properties of the NLIN and therefore it is suitable for addressing scenarios including fast adaptive nonlinear equalization [27], [28]. Furthermore, it is also compliant with four-dimensional modulation formats exhibiting inter-polarization dependency [29].

A fast and accurate method for performance assessments in nonlinear DWDM systems, named virtual lab (VL) has been proposed in [30] on the basis of the time-varying ISI model. It relies on separately generating the inter-channel NLIN according to its calculated second order statistics and adding it to the channel of interest (COI) which is split-step propagated through the fiber. This approach was shown to provide results with practically the same accuracy as of SSFTM but with enormous speedup enhancement factors. Yet, its main disadvantage is that it assumes perfectly linear modulation and ideal matched filtering, and hence it does not account for transmitter (TX) and receiver (RX) non-idealities that characterize practical systems.

Here, we report an extension of the VL, which we call the universal virtual lab (UVL) and whose preliminary version has been reported in [31]. The UVL allows accounting for non-ideal TX/RX implementations including bandwidth limitation, nonlinear device response (e.g., in the Mach-Zehnder modulator (MZM)), IQ imbalance, laser phase noise and link scenarios with wavelength selective switch (WSS) filters distributed along the optical link. The UVL also allows seamless addition of the NLIN to the COI samples in the optical domain while including Kerr effects and ISRS. We show that it provides a similar accuracy to SSFTM with a speedup factor exceeding 1000 with respect to conventional SSFTM in a broad range of system scenarios.

This paper is organized as follows. In Section II, we derive the underlying time-varying ISI model of the XPM contribution of the NLIN in the COI sample space and we present the principle of operation of the UVL. In Section III, we validate our results by comparing UVL and SSFTM simulations with different network scenarios (with and without ISRS) and two dual-polarization 16QAM modulation format variants with uniform and probabilistically shaped distribution, denoted DP-16QAM and DP-PS-16QAM respectively. The results include performance assessments in terms of pre-FEC Q-factor and generalized mutual information (GMI) after carrier phase recovery (CPR) and adaptive nonlinear equalization. In Section IV, we conduct performance assessments in an ultra-wideband (11 THz) C+L system and discuss equalization gain under different compensation scenarios. Finally, Section V is devoted to conclusions.

II. Principle of Operation

The VL approach [30] is based on the generation of noise which follows the analytically calculated second order statistics of the time-varying NLIN [22]. The VL applies the SSFTM to the COI only in order to account for chromatic dispersion, fiber loss and SPM effects while the nonlinear crosstalk induced by the ICs is added to the COI symbols at the receiver side after ideal matched filtering. This approach is highly efficient computationally and provides results that are indistinguishable from the SSFTM in a broad range of systems. The VL simulation runtime is nearly independent of the number of ICs since it depends mainly on the SSFTM of the COI and the pre-calculation of the XPM coefficients which is done in advance using efficient parallel computation. The shortcoming of the VL is that the evaluation of the quantities needed for correctly generating the NLIN relies on assumptions such as non-correlated symbols, linear modulation and ideal matched filtering. Therefore, the presence of RX imperfections and digital signal processing (DSP) performance limitations might lead to inaccurate generation of the NLIN that is added to the COI symbols. Furthermore, the NLIN impact on the RX DSP blocks, prior to COI symbol reconstruction such as the multiple input multiple output (MIMO) stage, cannot be evaluated with the VL.

The main idea for implementing the UVL is that rather than generating the NLIN based on analytically-obtained statistical properties (as in the VL [30]), in the UVL, we extract the statistical properties of the model numerically, by running a controlled subset of simple and fast two-channel simulations. These statistical properties are then used to generate the NLIN and allow its inclusion in the full system evaluation. The advantage of the UVL is that it enables seamless addition of the NLIN into the COI samples directly in the optical domain and allows accounting for non-ideal TX/RX scenarios and for the inclusion of WSS filters along the optical link.

In the following, variables written as bold-face letters refer to 2×1 vectors or to 2×2 matrices as can be deduced from the context. Non-bold variables refer to scalars.

A. Time-Varying ISI Model of the XPM Induced NLIN in the COI Sample Space

Within the time-varying ISI model and after ideal linear impairment compensation, the received COI symbols that are corrupted by the XPM and FWM induced NLIN are expressed as [30]

$$\begin{aligned}\hat{\boldsymbol{a}}_n &= \boldsymbol{a}_n + \Delta\boldsymbol{a}_n^{\mathrm{XPM}} + \Delta\boldsymbol{a}_n^{\mathrm{FWM}} \\ &= \boldsymbol{a}_n + i\sum_l \boldsymbol{R}_l^{(n)}\boldsymbol{a}_{n-l} + \Delta\boldsymbol{a}_n^{\mathrm{FWM}},\end{aligned} \tag{1}$$

where $\boldsymbol{a}_n = [a_{x,n},\, a_{y,n}]^T$ is a 2×1 vector whose elements are the n^{th} COI symbols in the two polarization tributaries and the term $\mathbf{R}_{\mathrm{l}}^{(\mathrm{n})}$ is the l^{th}-order 2×2 NLIN ISI matrix applied at time $\mathrm{n}\mathrm{T_s}$ with $\mathrm{T_s}$ denoting the symbol period. Equation (1) describes the XPM and FWM-induced NLIN contributions in the COI symbol space. In numerical simulation platforms, continuous time signal representation is performed with a discretized time grid at an appropriately large value of samples per symbol (SPS). Our objective is to extend the time-varying ISI model of the NLIN to the COI sample space in order to be able to add its contribution directly in the optical domain of the simulation platform in a seamless way, regardless of practical TX implementations. To do so, we notice that instead of adding the NLIN

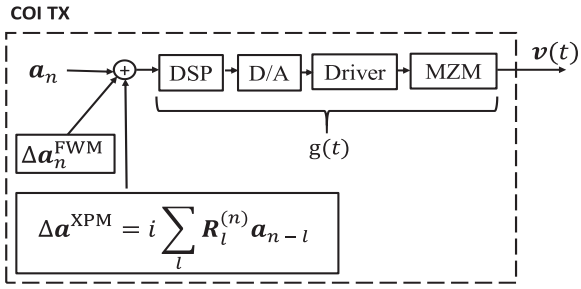

Fig. 1. NLIN model at the transmitter side. DSP: digital signal processor, D/A: digital to analog converter.

contribution to the COI symbols on the receiver side, we can equivalently add it on the transmitter side, as shown in Fig. 1. Indeed, within the first order perturbation analysis, XPM and FWM induced NLIN are not distorted by the COI induced SPM during fiber propagation.

In what follows, we will look at the signal samples at times $(n + k/N_{SPS})T_s$, where N_{SPS} is the number of samples per symbol, n counts the symbol durations and k is a time offset within the symbol duration ($k \in [0, 1, .., N_{SPS} - 1]$). For simplicity, the notation $X_{n+k/N_{SPS}}$ will be used instead of $X[(n + k/N_{SPS})T_s]$ for a generic time dependent function $X(t)$ and $\boldsymbol{X}_{n+k/N_{SPS}} \triangleq [X_{x,n+k/N_{SPS}}, X_{y,n+k/N_{SPS}}]^T$ will refer to a 2×1 vector whose elements are the waveform samples of the two polarization tributaries [2]. With this approach and under the assumption of a linear TX response, the COI samples at the TX output, $\boldsymbol{s}_{n+k/N_{SPS}}$ are accompanied by XPM-induced NLIN, $\Delta\boldsymbol{s}^{\mathrm{XPM}}_{n+k/N_{SPS}}$ and FWM-induced NLIN, $\Delta\boldsymbol{s}^{\mathrm{FWM}}_{n+k/N_{SPS}}$. Hence, they can be expressed as

$$\begin{aligned}\boldsymbol{v}_{n+k/N_{SPS}} &= \boldsymbol{s}_{n+k/N_{SPS}} + \Delta\boldsymbol{s}^{\mathrm{XPM}}_{n+k/N_{SPS}} + \Delta\boldsymbol{s}^{\mathrm{FWM}}_{n+k/N_{SPS}} \\ &= \sum_m \mathrm{g}_{n+k/N_{SPS}-m}\boldsymbol{a}_m \\ &\quad + \sum_m \mathrm{g}_{n+k/N_{SPS}-m}\Delta\boldsymbol{a}^{\mathrm{XPM}}_m + \Delta\boldsymbol{s}^{\mathrm{FWM}}_{n+k/N_{SPS}},\end{aligned} \tag{2}$$

where $\mathrm{g}_{n+k/N_{SPS}}$ is the TX impulse response which includes the pulse shaper, the digital pre-distortion, the digital to analog converter (D/A), the modulator driver and the optical modulator responses.

Using (1) and assuming that the TX response generates ISI of up to p symbols on each side of a given COI symbol, (2) can be rewritten in the form

$$\begin{aligned}\boldsymbol{v}_{n+k/N_{SPS}} &\approx \sum_{m=n-p}^{n+p} \mathrm{g}_{n+k/N_{SPS}-m}\boldsymbol{a}_m \\ &\quad + i\sum_{m=n-p}^{n+p} \mathrm{g}_{n+k/N_{SPS}-m} \sum_l \boldsymbol{R}_l^{(m)}\boldsymbol{a}_{m-l} + \Delta\boldsymbol{s}^{\mathrm{FWM}}_{n+k/N_{SPS}}.\end{aligned} \tag{3}$$

[2]Note that in Eq. (1) of [31], the time scale unit to be used with the fractional index convention should be $\mathrm{T_S}$, similarly to the present paper and not Δt, the sampling time unit, as erroneously stated there.

Assuming that the NLIN ISI matrix coefficients do not change noticeably within the effective temporal range of the memory distortion generated by the TX, we obtain the following approximation $\mathbf{R}_{\mathrm{l}}^{(\mathrm{m})} \approx \mathbf{R}_{\mathrm{l}}^{(\mathrm{n})}$ for $n - p \le m \le n + p$. As a result, in the second term of (3), the order of the two summation operators can be switched and after replacing the index m by $m' + l$, it can be shown that the COI samples corrupted by XPM and FWM induced NLIN can be expressed in the sample space in a similar way as expressed in (1)

$$\begin{aligned}\boldsymbol{v}_{n+k/N_{SPS}} &= \boldsymbol{s}_{n+k/N_{SPS}} + i\sum_l \boldsymbol{R}_l^{(n)}\boldsymbol{s}_{n+k/N_{SPS}-l} \\ &\quad + \Delta\boldsymbol{s}^{\mathrm{FWM}}_{n+k/N_{SPS}}.\end{aligned} \tag{4}$$

As noted earlier, the fact that instead of being added at the end of the link as in [30], the NLIN samples are added at the fiber input and then propagated with the COI is consistent with the first order perturbation framework. This facilitates the synchronization and alignment of the separately generated NLIN with the COI samples. The time-varying ISI model according to (1) and its extension in the COI sample-space (represented by (4)) have been derived under an assumption of linear modulation. TX modules usually exhibit nonlinear response originating in the D/A, the modulator driver and the optical modulator. In the relevant regime of moderate TX nonlinearity, it is reasonable to assume that only the signal samples $\boldsymbol{s}_{n+k/N_{SPS}}$ and the statistical properties of $\mathbf{R}_{\mathrm{l}}^{(\mathrm{n})}$ are affected by this TX nonlinearity, but once they are known, (4) remains valid. The accuracy of the UVL, which will be demonstrated numerically in Section III, testifies for the accuracy of this approximation.

B. UVL Methodological Flow

At the TX output, the COI samples corrupted by XPM and FWM induced NLIN contributions are expressed as

$$\begin{aligned}\boldsymbol{v}_{n+k/N_{SPS}} &= \boldsymbol{s}_{n+k/N_{SPS}} + \Delta\boldsymbol{s}^{\mathrm{XPM}}_{n+k/N_{SPS}} + \Delta\boldsymbol{s}^{\mathrm{XPM,residual}}_{n+k/N_{SPS}} \\ &\quad + \Delta\boldsymbol{s}^{\mathrm{FWM}}_{n+k/N_{SPS}},\end{aligned} \tag{5}$$

where $\Delta\boldsymbol{s}^{\mathrm{XPM}}_{n+k/N_{\mathrm{SPS}}} = i\sum_{l=-L}^{L} \boldsymbol{R}_l^{(n)}\boldsymbol{s}_{n+k/N_{\mathrm{SPS}}-l}$ from (4) is calculated up to the $+/-L^{th}$ ISI orders and $\Delta\boldsymbol{s}^{\mathrm{XPM,\ residual}}_{n+k/N_{\mathrm{SPS}}}$ accounts for the XPM contributed by higher orders of the ISI.

In order to generate these NLIN contributions, the UVL approach comprises two stages, as shown in Fig. 2 in the case of a WDM system with N+1 channels. In the first stage, the second order statistics (SOS) database which includes autocorrelation functions (ACFs) and pseudo-covariance functions (PCFs) of the NLIN ISI matrix elements, is generated up to the $+/-L^{th}$ ISI order. It uses a bank of fast SSFTMs based on two-channel coupled Manakov equations [18], each applied to the COI and one of the ICs, and an NLIN ISI matrix estimator as described in [34] (Further details are provided in sections II.C. and II.D). As the data of the different ICs are statistically independent, the database generation is suited for efficient parallel processing. Furthermore, since the SOS spectral evolution is slow at large frequency detuning, database decimation is possible and enables to reduce the SOS calculation to a subset of ICs and then to

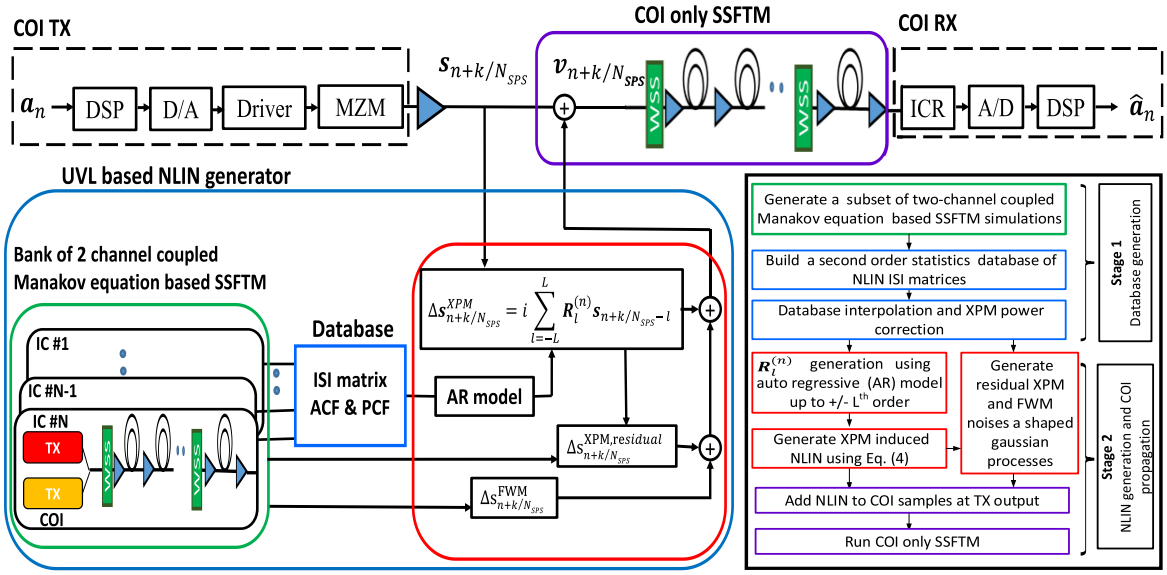

Fig. 2. UVL principle of operation. $\boldsymbol{a}_n$ and $\hat{\boldsymbol{a}}_n$ denote the 2 × 1 vector of transmitted and detected n$^{\text{th}}$ symbol, respectively. D/A: digital to analog converter, A/D: analog to digital converter, ICR: integrated coherent receiver, DSP: digital signal processor. Inset shows the UVL flow diagram.

proceed to database interpolation for the remaining ICs. The IC subset selection depends on the chromatic dispersion, the channel bandwidth and the channel spacing, since it is related to the walk-off between the COI and the IC pulses. ICs located at moderate frequency offset from the COI exhibit a SOS that evolves quickly and therefore they should be included in the database. With large frequency offsets, the ACFs and PCFs evolve slowly with the frequency offset within the symbol delay range of interest and consequently, can be decimated. The frequency offset threshold criterion beyond which IC subset decimation is possible, is provided in Appendix A.

In the second stage of the UVL, the ISI matrices, $\mathbf{R}_l^{(n)}$ are produced with the SOS evaluated in the first stage via an autoregressive (AR) model up to the $+/-L^{th}$ ISI orders as in [30] and then the XPM-induced NLIN is generated and added to the COI samples at the input of the optical link according to (4). The residual NLIN that is induced by higher ISI orders (beyond the ISI orders addressed by the AR model), as well as FWM noise (whose power is calculated using [32] according to the EGN model) are added as shaped Gaussian noise sources to the COI samples. The COI signal with the added NLIN is then split-step propagated through the link in order to accurately account for single-channel propagation effects[3].

A key advantage of the UVL is in the fact that when ISRS is negligible, the channel power profile is only scaled by a scalar factor when varying the optical launch power levels. Therefore, the SOS database needs to be generated only once for a given TX and physical link scenario with a single moderate optical power level. Then, the database can be trivially scaled to a desirable power level without rerunning the entire computation [30]. This allows rapid assessment of the NLIN system impact for arbitrary power levels and for different receiver settings. When ISRS is not negligible, the SOS database needs to be generated for each optical launch power profile. In addition, the FWM power calculation takes into account the ISRS induced power profile of the channels.

[3]It should be noted that both with VL and with UVL, one can increase the accuracy of the simulation by propagating the COI with a desired number of its nearest neighboring channels using SSFTM and adding the remaining interference semi-analytically. However, this will result in an increase of the simulation runtime.

C. Two-Channel Coupled Manakov Equations

Since the purpose of the set of two-channel Manakov equations based SSFTM is only to estimate XPM-induced NLIN parameters, we can proceed to further simplifications for speedup enhancement. First, we turn off the SPM effect on the COI. Secondly, within the first order perturbation analysis, the IC undergoes linear effects only while the COI is affected by both linear and XPM effects. As a results, the set of two-channel coupled Manakov equation from [18] is simplified as follows

$$\begin{cases} \frac{\partial}{\partial z}\boldsymbol{s}_1(z,t) = \left(\hat{D}_1 + \hat{\boldsymbol{N}}_1\right)\boldsymbol{s}_1(z,t), \\ \frac{\partial}{\partial z}\boldsymbol{s}_2(z,t) = \hat{D}_2\boldsymbol{s}_2(z,t), \end{cases} \tag{6}$$

where $\boldsymbol{s}_i(\mathrm{z},\mathrm{t}) \triangleq [s_{x,i}(z,t), s_{y,i}(z,t)]^T$is the two-dimensional complex envelop Jones vector and the index $i \in \{1,2\}$ refers to the COI and IC, respectively.

The linear operators are expressed as

$$\begin{aligned} \hat{D}_1 &= -\frac{\alpha_1(z)}{2} + j\frac{\beta_{2,1}}{2}\frac{\partial^2}{\partial t^2}, \\ \hat{D}_2 &= -\frac{\alpha_2(z)}{2} + d\frac{\partial}{\partial t} + j\frac{\beta_{2,2}}{2}\frac{\partial^2}{\partial t^2}, \end{aligned} \tag{7}$$

where α_i, $\beta_{2,\mathrm{i}}$ are the effective fiber loss, the second order dispersion coefficients at the COI ($i = 1$) and IC ($i = 2$) frequencies, respectively and d refers to the walk-off parameter between the IC and the COI and it includes the contribution of the dispersion slope.

The effective fiber loss coefficient is frequency and distance dependent since it includes the intrinsic fiber loss parameter, $\alpha_{\text{fiber}}(f_i)$ and the ISRS induced excess gain/loss coefficient, $\alpha_{\text{ISRS}}(f_i, z)$, both applied at the channel frequency f_i as shown below

$$\alpha_i(z) = \alpha_{\text{fiber}}(f_i) + \alpha_{\text{ISRS}}(f_i, z), \tag{8}$$

where $\alpha_{\text{ISRS}}(f_i, z)$ is derived by solving the set of N_{ch} coupled power differential equations describing the power exchange due to ISRS with N_{ch} being the total number of channels (NOC) [11] as follows

$$\frac{\partial P_i}{\partial z} = -\sum_{k=i+1}^{N_{ch}} \frac{f_i}{f_k} g_r(\Delta f) P_k P_i + \sum_{k=1}^{i-1} g_r(\Delta f) P_k P_i - \alpha_{\text{fiber}}(f_i) P_i, \tag{9}$$

where $g_r(\Delta f)$ is the normalized simulated Raman scattering (SRS) gain at frequency offset $\Delta f = f_i - f_k$, P_i and P_k are the powers of the channels i and k, with $k = 1$ denoting the highest frequency.

The nonlinear XPM operator applied to the COI is expressed as [18]

$$\hat{\boldsymbol{N}}_1 = -j\frac{8}{18}\gamma\left(3|\boldsymbol{s}_2(z,t)|^2 + \boldsymbol{S}_2^{\text{Stokes}}(z,t)\cdot\boldsymbol{\sigma}\right), \tag{10}$$

where γ stands for the fiber nonlinear coefficient, $\boldsymbol{\sigma}$ is a vector of Pauli matrices [33] and $\boldsymbol{S}_2^{\text{Stokes}} = \boldsymbol{s}_2^{\dagger}(z,t)\boldsymbol{\sigma}\boldsymbol{s}_2(z,t)$ is the Stokes vector representation of the IC Jones vector.

D. ISI Matrix Estimation in the Sample Space

For each IC, we extract the SOS of the ISI matrices resulting from the two-channel coupled Manakov equation based SSFTM simulation. At the link end, according to (4), the ISI matrices are stationary within the SPS phase offset range $k \in [01, .., N_{\text{SPS}} - 1]$ and each ISI matrix estimation can be obtained at each SPS phase offset in a similar way as reported in [34]

$$\hat{\boldsymbol{R}}_l^{(n,k)} = -i\left(\boldsymbol{v}_{n+k/N_{\text{SPS}}} - \boldsymbol{s}_{n+k/N_{\text{SPS}}}\right)\boldsymbol{b}_{n+k/N_{\text{SPS}}-l}^T, \tag{11}$$

where the COI samples $\boldsymbol{v}_n$ are obtained at the end of the link after CD compensation in the optical domain and $\mathbf{b}_n$ is defined as

$$\boldsymbol{b}_n = \left[s_x^{-1}(nT_S), s_y^{-1}(nT_S)\right]^T, \tag{12}$$

Afterwards, these ISI matrix estimations are averaged over the SPS phase offset range in order to increase their accuracy.

$$\hat{\boldsymbol{R}}_l^{(n)} = \frac{1}{N_{\text{SPS}}}\sum_{k=0}^{N_{\text{SPS}}-1}\hat{\boldsymbol{R}}_l^{(n,k)}, \tag{13}$$

Using (13), we calculate the ACFs and PCFs of the ISI matrix components with time difference Δn and order l that will be used to build the AR model,

$$C_l^{ij}(\Delta n) = E\left(R_{l,ij}^{(n)} R_{l,ij}^{*(n+\Delta n)}\right), \tag{14}$$

$$D_l^{ij}(\Delta n) = E\left(R_{l,ij}^{(n)} R_{l,ij}^{(n+\Delta n)}\right), \tag{15}$$

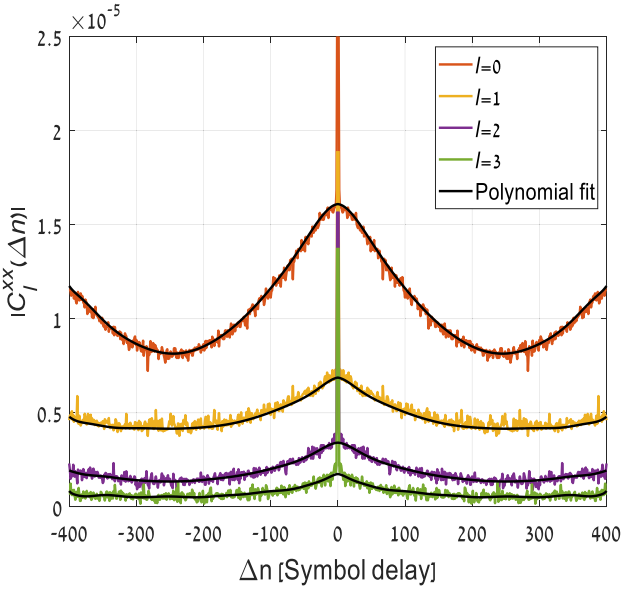

Fig. 3. Absolute value of the ACF of the $\mathrm{R}_{1,\mathrm{xx}}^{(\mathrm{n})}$ ISI component for $l = 01{,}23$, extracted from the two-channel Manakov equation based SSFTM with frequency offset of 525 GHz and optical launch power of 0 dBm. Black lines are obtained by the polynomial fitting of the ACFs.

where * denotes complex conjugate and $\mathrm{R}_{1,\mathrm{ij}}^{(\mathrm{n})}$ is the ij component of the ISI matrix $\mathbf{R}_1^{(\mathrm{n})}$ with $i, j \in \{x, y\}$, referring to the x or y polarization tributaries.

Since the ISI matrices are characterized by a moving average model, their corresponding covariance and pseudo-covariance matrices are smooth while the numerical estimations of Eqs. (14), (15) are noisy. In order to improve the accuracy and stability of the AR model, we apply polynomial fitting to the ACF and PCF estimates as a denoising technique. Furthermore, the estimation error of the ACF and PCFs exhibits a delta correlated term around $\Delta\mathrm{n} = 0$ that can be easily removed via interpolation due to the smooth nature of the ISI matrix components [34]. Fig. 3 shows the absolute value of the calculated ACF of the $\mathrm{R}_{1,\mathrm{xx}}^{(\mathrm{n})}$ ISI component for $l = 0, 1, 2, 3$ using test case I from Table I with $N_{SPS} = 5$. The ACF is estimated from the two-channel Manakov equation based SSFTM with a frequency offset of 525 GHz between the IC and the COI and an optical launch power of 0 dBm. As expected, the strength of the ACF decreases with the ISI order. The ACFs of the negative ISI orders are identical to their ACF counterpart with positive ISI orders (not shown in Fig. 3 for clarity purposes). The delta correlated term is present in all the ISI orders and removed by polynomial fitting.

The interpolated value of the autocorrelation functions at $\Delta\mathrm{n} = 0$ allows estimation of the IC induced XPM contribution to the NLIN power for each l^{th} ISI order on both polarizations, noted $\mathrm{P}_l^{\mathrm{XPM},IC}$. In the case of independent symbol data on the two polarization tributaries, we obtain

$$P_l^{\mathrm{XPM,IC}} = C_l^{xx}(0) P_{\mathrm{COI}}^x + C_l^{xy}(0) P_{\mathrm{COI}}^y + C_l^{yx}(0) P_{\mathrm{COI}}^x + C_l^{yy}(0) P_{\mathrm{COI}}^y, \tag{16}$$

where $\mathrm{P}_{\mathrm{COI}}^{x/y}$ is the COI optical power at polarization x/y.

TABLE I
TEST SCENARIOS

Test scenario	I	II
Fiber type	G.652	G.652
Span number	15	25
Span length [km]	100	100
Dispersion [ps/nm-km] @1550 nm	17	17
Dispersion slope [ps/nm^2-km] @1550 nm	0.06	0.06
Nonlinear coefficient [1/W/km]	1.3	1.3
Fiber Loss [dB/km]	0.2	0.2
ISRS	no	yes*
Modulation format	DP-16QAM	DP-PS-16QAM
Symbol entropy [bit/4D-symb]	8	5.6
Symbol Rate [GBd]	68.8	120
Channel Spacing [GHz]	75	125
Number of channels	3-45	21
Roll-off factor	0.1	0.1
TX/RX laser linewidth [kHz]	100	100
TX/RX bandwidth [GHz]	31/31	50/44
TX digital pre-distortion	yes	yes
Nonlinear MZM response	yes	yes
TX IQ imbalance [degree]	+/- 5	+/- 5
WSS filter bandwidth [GHz]	72	122
Number of cascaded WSS filters	8	10
Number of symbols per channel (full-bandwidth SSFTM and COI-only SSFTM)	193248	
Number of symbols per channel (two-channel coupled Manakov equations)	96624	

*Compressed ISRS model

E. XPM Power Correction

The total XPM power applied to the COI is the result of all the ISI order power contributions from all the ICs. Since the AR model generates the NLIN up to the $+/-L^{th}$ ISI order, it is necessary to estimate the residual XPM noise that will be added as a shaped Gaussian noise. The residual XPM power is the sum of the contribution of all the residual XPM powers from the ICs and it is expressed as

$$P_{\text{residual}}^{\text{XPM}} = \sum_{i=1}^{N_{IC}} P^{\text{XPM},i} - P_L^{\text{XPM},i} = \sum_{i=1}^{N_{IC}} \sum_{l=-\infty}^{+\infty} P_l^{\text{XPM},i} - \sum_{i=1}^{N_{IC}} \sum_{l=-L}^{+L} P_l^{\text{XPM},i}, \quad (17)$$

where N_{IC} is the total number of ICs, $\text{P}^{\text{XPM,i}}$ is the total XPM power contribution of the i^{th} IC, (which includes all the NLIN ISI orders) extracted from the two-channel Manakov equation based SSFTM and $\text{P}_L^{\text{XPM,i}} = \sum_{l=-L}^{+L} \text{P}_l^{\text{XPM,i}}$ is the sum of all XPM power contribution from the i^{th} IC, up to the $+/-L^{th}$ ISI order, calculated according to (16).

The estimation of $\text{P}_{\text{residual}}^{\text{XPM}}$ in (17) relies on the accurate extraction of $\text{P}^{\text{XPM,i}}$ from the numerical solution of the coupled two-channel nonlinear equations. This extraction turns out to be very noisy and strongly dependent on the step-size used in the numerical solution. This issue is illustrated in Fig. 4(a), which shows the value of $\text{P}^{\text{XPM,i}}$ as a function of detuning between the considered IC and COI for two different step sizes. The figure is plotted with the parameters of test-case I of Table I with 41 channels (DWDM bandwidth of 3.075 THz), each one having an optical launch power of 0 dBm and with COI located at the high-frequency edge of the DWDM spectrum. As it is evident in the figure, the larger step-size (represented by nonlinear phase offset of $\Delta\,\phi_{\text{NL}}^{\text{Max}} = 0.1^{\circ}$) results in an artificial floor at large frequency detunings [35] while the XPM power should reduce properly with the frequency detuning between the COI and IC [36]. This floor does not exist when the step-size is reduced ($\Delta\,\phi_{\text{NL}}^{\text{Max}} = 0.01^{\circ}$). This implies that in order to avoid significantly overestimating $\text{P}_{\text{residual}}^{\text{XPM}}$, the step-size in the two-channel simulations should be very low.

The need to reduce the step-size with the frequency offset in the two-channel simulations imposes an undesirable constraint and severely slows down the procedure of evaluating the XPM power. In order to circumvent this difficulty, we extract $\text{P}_l^{\text{XPM,i}}$ from (16), but instead of evaluating $\text{C}_l^{pq}(\Delta\text{n})$ (where p and q each assume the values x and y) numerically for $\Delta\text{n} = 0$, as written in (16), we evaluate them for nonzero values of Δn and make use of the smoothness of the ACF in order to estimate $\text{C}_l^{pq}(0)$, consistently with our discussion in the context of Fig. 3. Fig. 4(b) demonstrates that the estimation of the XPM power order contributions for each IC, $\text{P}_l^{\text{XPM,i}}$ is similar for both step sizes. As a consequence, $\text{P}_L^{\text{XPM,i}}$ can be used to estimate the correct slope of the XPM power contribution per IC as function of the frequency offset, leading to a more accurate estimation of total XPM power. As shown by the dotted curve in Fig. 4(a), after XPM power correction with $\Delta\phi_{\text{NL}}^{\text{Max}} = 0.1^{\circ}$, the estimation of the evolution of $P^{\text{XPM,i}}$ with the frequency detuning is now perfectly consistent with the curves that are obtained with $\Delta\,\phi_{\text{NL}}^{\text{Max}} = 0.01^{\circ}$.

III. NUMERICAL VALIDATION

In order to evaluate the accuracy of the UVL method with respect to the standard full–bandwidth SSFTM, pre-FEC Q-factor and GMI assessments are performed for the different link scenarios as described in Table I. Two modulation formats are considered: 68.8 GBd DP-16QAM and 120 GBd DP-PS-16QAM with a symbol entropy of 5.6 bit/4D-symb. The COI and IC signals are distorted from the TX module due to digital pre-emphasis, TX limited bandwidth, nonlinear MZM response, and IQ imbalance of up to $+/-5$ degrees. In both scenarios, the digital pre-distortion along with the TX frequency response lead to a memory generated distortion in the order of $+/-15$ symbol intervals which remains below the correlation length of the ISI matrices $\mathbf{R}_{\text{l}}^{(\text{n})}$ for at least $|l| \leq 3$, ensuring the validity of (4)[4]. A laser linewidth of 100 kHz is used on both TX and RX sides. G.652 fiber type is considered with standard values of dispersion, nonlinear fiber coefficient and fiber loss. Polarization mode dispersion and polarization dependent loss are not included in the simulation. The WSS filters are distributed

[4]The correlation length is defined as the symbol delay leading to an amplitude reduction by 50% of the ACFs. In test scenario I, when the number of channel is 15 or higher, the correlation length of the ISI matrices is higher than 15 symbol intervals for $|l| \leq 7$. In test scenario II with 21 channels, the correlation length is also higher than 15 symbol intervals for $|l| \leq 7$.

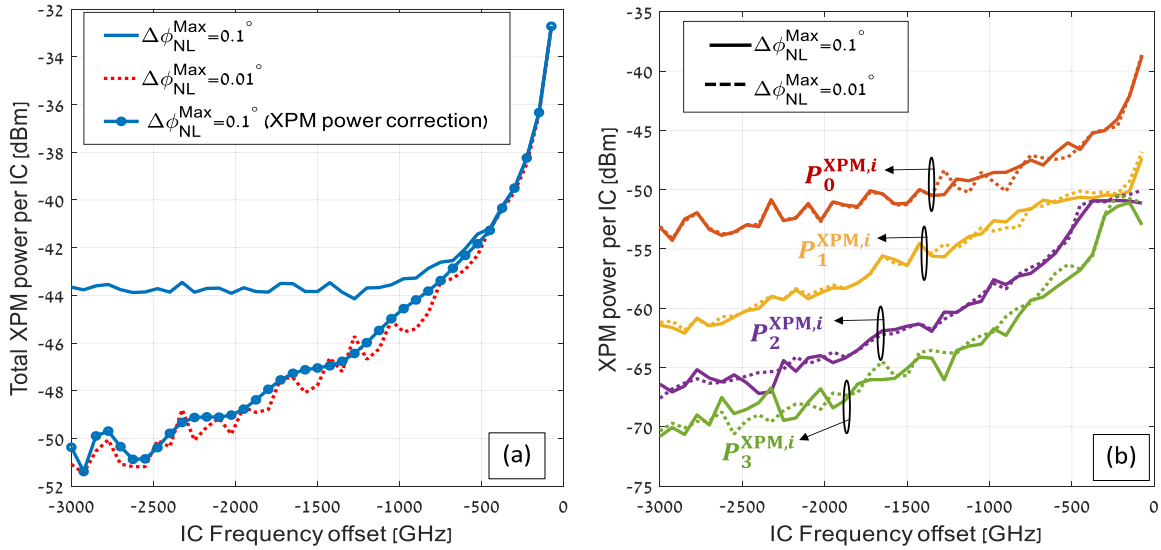

Fig. 4. (a) XPM power contributed by IC as a function of the frequency offset with respect to the COI, as obtained from the set of two-channel simulations. Results are shown for two step-sizes $\Delta\phi_{\mathrm{NL}}^{\mathrm{Max}} = 0.1°$ and $\Delta\phi_{\mathrm{NL}}^{\mathrm{Max}} = 0.01°$. In a brute force computation, the larger step-size results in an artificial floor (thin blue curve) at high detunings. This floor is eliminated when the proposed computation is performed (blue curve with circles). (b) Same as in Fig. (a) except that the contribution of the first 4 ISI orders are displayed separately. Here, no floor is observed because the contribution of the individual orders is evaluated as discussed in part D of this section.

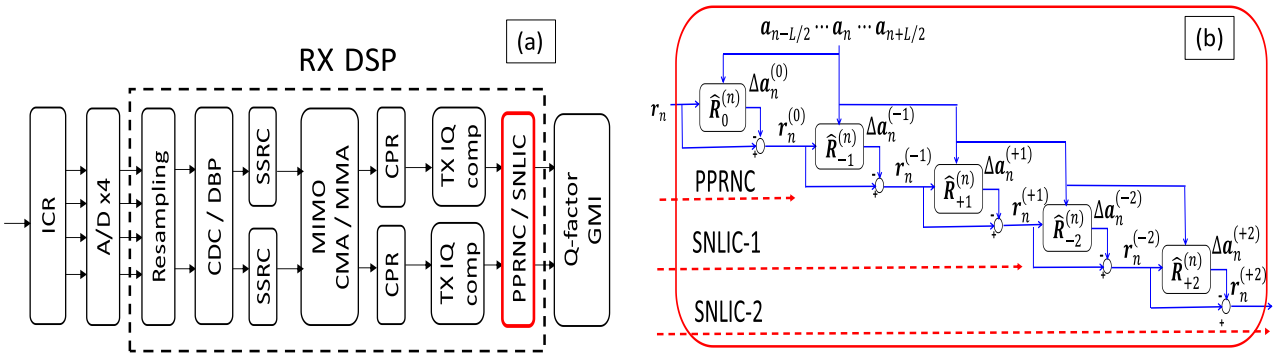

Fig. 5. Coherent receiver block diagram with DSP block chain (a), Genie mode based PPRNC and SNLIC schemes up to the $+/-2^{\mathrm{nd}}$ ISI order (b). ICR = Integrated coherent receiver, A\D = Analog to digital converter.

along the links with a period of every 2 spans for scenario I and every 3 spans for scenario II. We use lumped amplification that ideally compensates for the span loss, and the amplified spontaneous emission (ASE) noise is loaded at the RX input.

In the UVL scheme, the NLIN SOS database is generated up to $+/-7^{\mathrm{th}}$ ISI order (as defined in (4)), using a bank of two-channels fast SSFMs for a subset of ICs and then interpolated for the remaining ICs. The IC subset selection depends on the chromatic dispersion, the channel bandwidth and the channel spacing. For the test scenarios from table I, the database decimation starts for IC frequency offset larger than 500 GHz with IC selected every 450 GHz and 500 GHz for test scenarios I and II, respectively. The split-step simulation of the COI is performed with 5 samples per symbol. The generated NLIN is added to the COI samples at the TX output and after proper power scaling. In order to demonstrate the UVL accuracy, the entire system is also simulated using the full-bandwidth SSFTM (with the SPS being dictated by the number of channels). A total of 96624 symbols per channel are used for the SOS database generation and 193248 symbols are used for simulating the propagation of the COI within the UVL, as well as in the full-bandwidth SSFTM simulations.

As shown in Fig. 5(a), after coherent detection, ideal digital back propagation (DBP), or chromatic dispersion compensation (CDC) are applied, followed by square root raised cosine (SRRC) filtering and adaptive compensation techniques such as MIMO and CPR. MIMO is implemented via the constant modulus algorithm and the multiple modulus algorithm (CMA/MMA) while the CPR is implemented with blind phase search (BPS) and maximum likelihood estimator (MLE) methods [37]. This is followed by transmitter IQ compensation based on the Gram Schmidt orthogonalization procedure [38]. In order to demonstrate the power and accuracy of the UVL method, we also consider cases in which nonlinear compensation methods are applied. As we mentioned earlier, one of the major advantages of the UVL relies on its ability to accurately model the

time dynamics of the NLIN which allows precise performance predictions when using adaptive equalization schemes. These methods include "genie-mode" based phase and polarization rotation noise compensation (PPRNC) as described in [39] and its extension to the nonlinear compensation of the first and second-orders of NLIN using a successive nonlinear interference cancellation scheme (SNLIC-1 and SNLIC-2 respectively) as shown in Fig. 5(b). By genie mode, we assume that the true transmitted symbols $\mathbf{a}_{\mathrm{n}}$ are used on the RX side for the estimation and the compensation of the different NLIN orders. The purpose of the genie mode regime is to allow an accurate comparison of the NLIN statistics up to the 2nd ISI order, generated by the UVL and the full-bandwidth SSFTM with respect to the transmission system performances. As it will be discussed in Section IV, realistic PPRNC and SNLIC implementations in a non-genie mode (by using hard decoded symbols on the RX side) can provide a compensation gain up to the 2nd ISI order in a high BER regime. Ultimately, the genie mode applied to PPRNC and SNLIC methods can be seen as their performance limit in a turbo equalization architecture, when the FEC decoder operates in an error-free regime. In the SNLIC method, the XPM contributions of the different NLIN orders are removed iteratively from the noisy symbols starting from the 0th order as shown below

$$\boldsymbol{r}_n^{p(k)} = \boldsymbol{r}_n^{p(k-1)} - \Delta\boldsymbol{a}_n^{p(k)} = \boldsymbol{r}_n^{p(k-1)} - i\hat{\boldsymbol{R}}_{p(k)}^{(n)}\boldsymbol{a}_{n-p(k)}, \quad (18)$$

where $p = [0, -1, 1, -2, 2, \ldots]$ refers to the ordered sequence of NLIN orders and $k = [0,\ 1,\ 2, \ldots]$ denotes the iteration number. For the case $k = 0$, we get

$$\boldsymbol{r}_n^{(0)} = \boldsymbol{r}_n - i\hat{\boldsymbol{R}}_0^{(n)}\boldsymbol{a}_n, \quad (19)$$

where $\mathbf{r}_{\mathrm{n}}$ refers to the 2×1 noisy symbol vector after the TX IQ compensation stage. The 0th order least square ISI matrix estimation derived in [39] is extended to any ISI order as follows

$$\hat{\boldsymbol{R}}_{p(k)}^{(n)\dagger} = i\left(\sum_{j=n-L/2}^{n+L/2} \boldsymbol{a}_{j-p(k)}\boldsymbol{a}_{j-p(k)}^{\dagger}\right)^{-1} \times \left(\sum_{j=n-L/2}^{n+L/2} \boldsymbol{a}_{j-p(k)}\left(\boldsymbol{r}_j^{p(k-1)} - \boldsymbol{a}_j\right)^{\dagger}\right), \quad (20)$$

where L+1 is the window averaging length and † is the transpose conjugate operator[5]. For test scenarios I and II, a window length of 101 symbols is used. We also compare the computational runtime of the UVL and SSFTM by defining the speedup factor (SUF) as the ratio between the full-bandwidth SSFTM and UVL computational runtimes, noted $T_{\text{full-BW-SSFTM}}$ and T_{UVL}, respectively

$$SUF = \frac{T_{\text{full-BW-SSFTM}}}{T_{\mathrm{UVL}}} = \frac{T_{\text{full-BW-SSFTM}}}{T_{\mathrm{DB}} + T_{\mathrm{NLIN}} + T_{\text{COI-SSFTM}}}, \quad (21)$$

[5]The PPRN estimation $\hat{\boldsymbol{H}}_0^{(n)}$ from Eq. (31) from [39] and $\hat{\boldsymbol{R}}_0^{(n)}$ are related as $\hat{\boldsymbol{H}}_0^{(n)} = \boldsymbol{I} + i\hat{\boldsymbol{R}}_0^{(n)}$ with $\boldsymbol{I}$ being the 2×2 identity matrix.

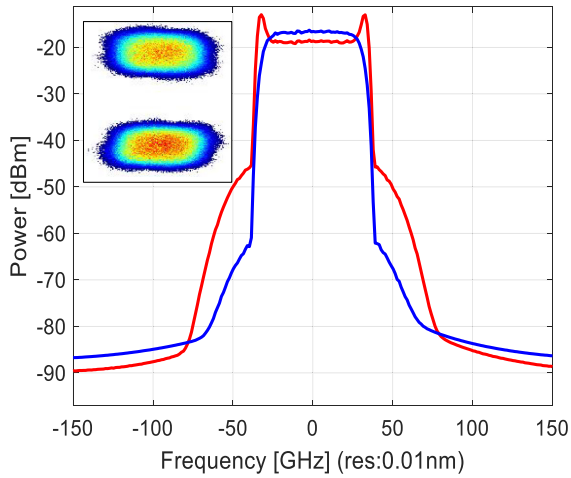

Fig. 6. COI optical spectrum at the transmitter output without digital pre-distortion and with linearized MZM response (blue) and with digital pre-distortion and nonlinear MZM response (red). The inset shows the COI sample diagram at the transmitter output corresponding to the former case (red).

where T_{DB}, T_{NLIN}, $T_{\text{COI-SSFTM}}$ represent the computational runtime required for the database generation, the NLIN generation and the single channel SSFTM, respectively. As the XPM noise contribution of the different ICs are statistically independent, the database generation is compliant with full parallel processing. The computation runtime required to generate the NLIN includes the generation of both XPM and FWM induced NLIN.

A. Test Scenario I: 68.8 GBd DP-16QAM

For test scenario I, we consider 68.8 GBd DP-16QAM channels over 15 spans of 100 km of standard single mode fiber (G.652) with a non-ideal TX response. ISRS is negligible since in the investigated test cases, the ISRS induced power tilt between the edge channels is below 0.2 dB.

Fig. 6 shows the COI optical spectrum at the transmitter output exhibiting spectral broadening due to the MZM nonlinear response and "bat ears" like digital pre-emphasis used to partially compensate the WSS spectral narrowing. The COI is placed at the central channel of the DWDM system and the OSNR is set to 22 dB for a launch power of 3 dBm. Ideal digital back-propagation (DBP) is applied to the COI in order to fully focus on XPM and FWM induced NLIN.

In Fig. 7(a), we show the estimated pre-FEC Q-factor and GMI after DBP and SNLIC based nonlinear compensation method up to the second ISI order. The number of channels (NOC) varies between 3 and 45, and the per-channel launch power is set to 3 dBm. The excellent agreement between the full-bandwidth SSFTM and the UVL is self-evident in all cases with the differences remaining below the level of 0.05 dBQ and 0.02 bit/4D-symb for Q-factor and GMI, respectively. Fig. 7(b) shows the speedup factor of the UVL method with respect to the full-bandwidth SSFT runtime for the cases of 15, 25 and 45 channels as a function of the number of parallel processes used to generate the SOS database. In this case, the UVL runtime includes the database generation and the COI split-step propagation. As noted

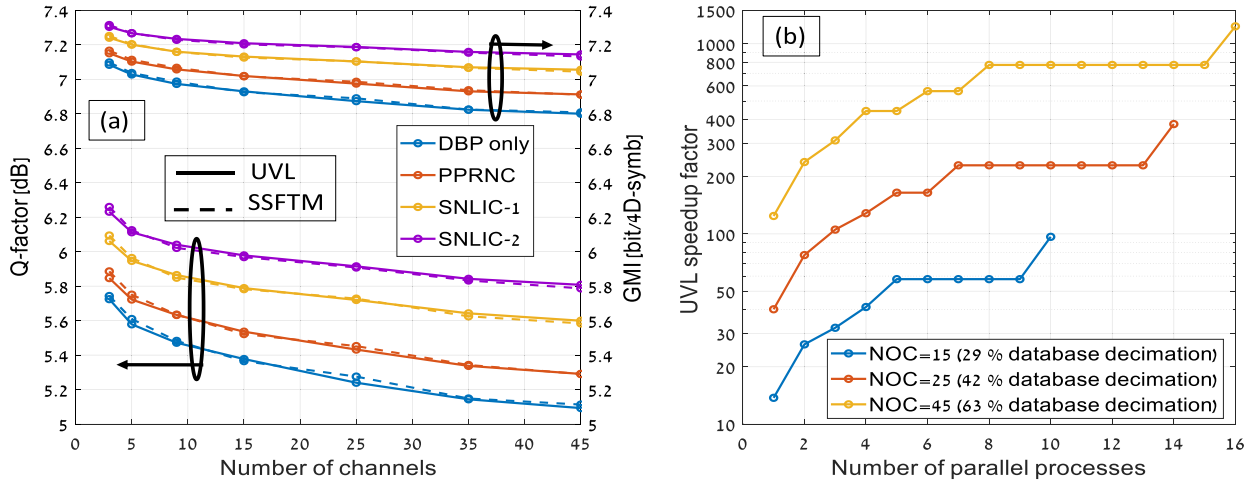

Fig. 7. Test scenario I: (a) Q-factor and GMI as function of NOC for UVL (solid lines) and SSFTM (dashed lines) after DBP and SNLIC, (b) SUF as a function of the number of parallel processes for NOC=15, 25 and 45. For both figures, the optical channel launch power is 3 dBm.

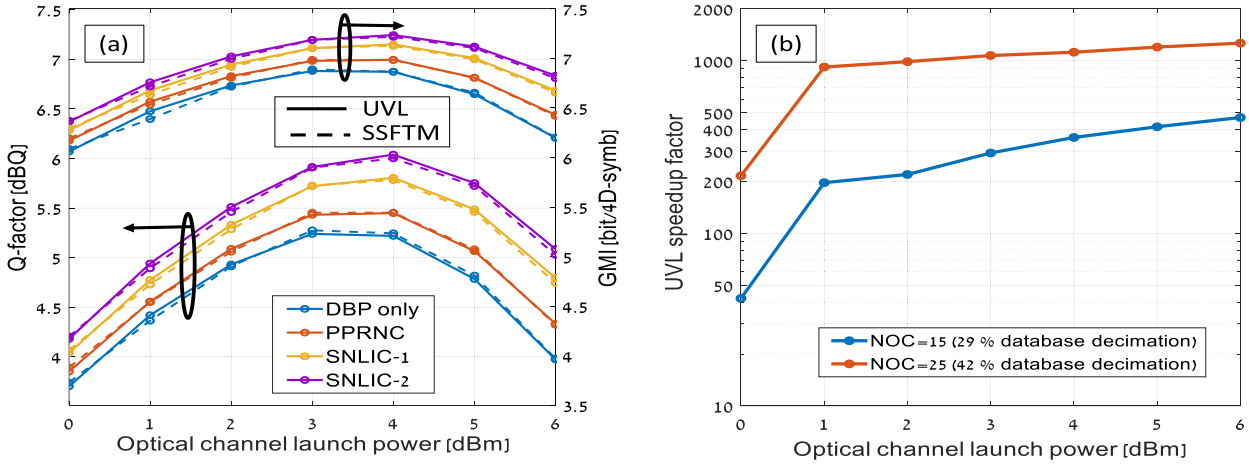

Fig. 8. Test scenario I: (a) Q-factor and GMI as function of the optical channel launch power for UVL (solid lines) and SSFTM (dashed lines) with NOC=25 after DBP and SNLIC, (b) SUF as a function of the optical launch power for NOC=15 and 25.

earlier, not all ICs are evaluated when collecting the NLIN SOS. We define the database decimation factor as the ratio between the interpolated IC number and the total IC number. The decimation factor increases with the NOC, reaching 29%, 42% and 63% for 15, 25 and 45 channels, respectively. The UVL SUF increases with the NOC and the number of parallel processes. It ranges from 96.5 to over 1388 in the tested cases when using full parallel processing capabilities. The SUF is expected to be considerably larger in the case of fully loaded ultra-wideband systems where the full-bandwidth SSFTM runtime becomes totally unrealistic.

Another advantage of the UVL is in the power scaling of the SOS database that can be performed analytically [30], as long as ISRS is negligible. This aspect is explored in Fig. 8(a), where the pre-FEC Q factor and the GMI are shown as a function of the per-channel launch power with a 25-channel system. The SOS database is generated for a reference per-channel launch power of 0 dBm and power-scaled for the other power levels. The difference between the results of the UVL and the SSFTM is smaller than 0.1 dBQ and 0.08 bit/4D-symb in all cases for the Q-factor and GMI, respectively. Similar accuracy is obtained for NOC = 15 (not shown in the figure for clarity purposes). Fig. 8(b) shows the SUF as a function of the optical per-channel launch power with 15 and 25 channels. The SUF takes into account the database generation with the maximum number of parallel processes for the reference launch power level while the database is power-scaled for the other power levels. Clearly, the SUF increases sharply with the database power scaling and the NOC.

B. Scenario Test II: 120 GBd DP-PS-16QAM With ISRS

For test scenario II, we consider 21 × 120 GBd DP-PS-16QAM channels with a symbol entropy of 5.6 bit/4D-symb. The transmission system includes 25 spans of 100 km of standard single-mode fibers (G.652) with non-ideal TX response including digital pre-emphasis and nonlinear MZM response. In order to verify the UVL compliance with the ISRS effect, we use a fiber model with a normalized stimulated Raman scattering

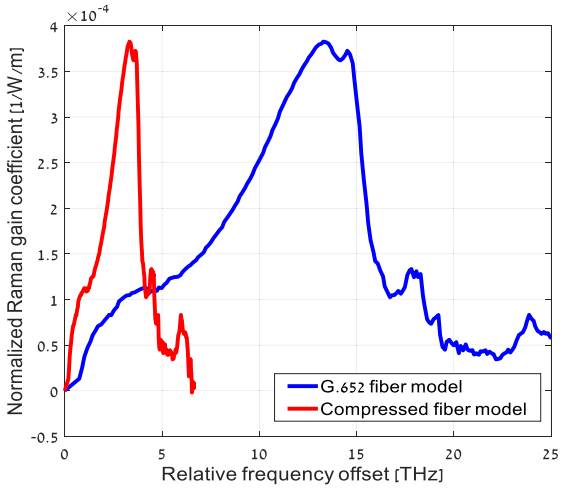

Fig. 9. Normalized Raman gain coefficient with respect to the relative frequency offset for G.652 fiber (blue) and compressed Raman fiber model by a factor of 4 (red).

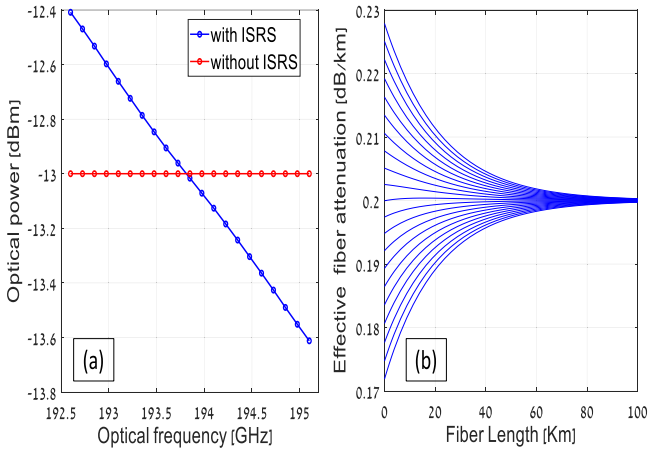

Fig. 10. Test scenario II with 21 channels having a uniform launch power profile of 7 dBm. (a) Optical power profile after transmission over 100 km without ISRS (red) and with compressed ISRS (blue). (b) Effective fiber loss coefficient for each transmitted channel as function of the fiber length.

(SRS) gain spectrum compressed by a factor of 4 as shown in Fig. 9. This prevents from running SSFTM simulations with a large number of channels that would be impractical in terms of computational runtime. We consider the case where all the channels are launched with a uniform optical power and there is an ideal optical power equalization at every span. Fig. 10(a) shows the spectral power profile after transmission of 100 km with a launch channel power of 7 dBm. The compressed SRS gain model leads to a power tilt of 1.2 dB over a DWDM bandwidth of 2.625 THz. Fig. 10(b) shows the evolution of the effective fiber loss coefficient for each channel along one fiber span of 100 km, according to (8). Throughout the ISRS process, higher frequency channels are depleted via partial power transfer towards lower frequency channels. This results in an effective fiber loss parameter that is lower than the intrinsic fiber loss parameter for lower frequency channels while it increases beyond the intrinsic fiber loss parameter for higher frequency channels.

In Fig. 11, the pre-FEC Q-factor and the GMI of the COI located at the central channel are shown when varying the optical launch power from 2 dBm to 10 dBm with a 21-channel system for the ideal DBP case (Fig. 11(a)) and the CDC case (Fig. 11(b)). The OSNR is set to 23 dB at launch power of 7 dBm. Since ISRS is non-negligible in this test scenario, each channel power profile changes with the optical launch power level. As a consequence, the SOS database is generated for every launch power level but for 12 ICs out of 20 ICs while the database is interpolated for the remaining 8 ICs (resulting in a decimation factor of 40%). The difference between the results of the UVL and the SSFTM is smaller than 0.1 dBQ and 0.09 bit/4D-symb in all cases for the Q-factor and GMI, respectively. Fig. 11(c) shows the SUF as a function of the optical per-channel launch power and the contribution of the database generation in the total UVL runtime. The SUF takes into account the database generation for each optical launch power with a maximum number of 12 parallel processes. Again, the SUF increases with the optical launch power, ranging from 153 to 525 with a SUF of 375 at optimum launch power of 7 dBm. The contribution of the database generation runtime decreases with the optical launch power as the COI split step propagation runtime increases. The sharp SUF behavior observed with the optical launch power in test scenarios I (see Fig. 7(b)) is not present here since the database is generated for each optical launch power case and represents at least 48.9% of the total UVL runtime.

Fig. 12 shows the pre-FEC Q-factor and the GMI assessments for all 21 channels in the case of a per-channel launch power of 7 dBm with a fixed OSNR of 23 dB in the cases of ideal DBP and SNLIC (Fig. 12(a)) and CDC and SNLIC (Fig. 12(b)). In both cases, by applying the UVL to different COI optical frequencies within the DWDM spectrum, the accuracy remains within 0.1 dBQ and 0.04 bit/4D-symb for Q-factor and GMI, respectively.

IV. Performance Assessments in an Ultra-Wideband C+L DWDM System

The numerical validation presented in the previous section enables the use of the UVL as a very fast and accurate simulation tool in ultra-wideband scenarios where the full-bandwidth SSFTM is impractical. Here, we conduct a performance analysis for an ultra-wideband transmission system with 88 × 120 GBd DP-PS-16QAM channels, accounting for a total DWDM bandwidth of 11 THz with a continuous occupation of the C and L bands. We use an optical link similar to the one from test scenario II with the exception that the compressed normalized SRS gain coefficient model is replaced by the conventional normalized SRS gain coefficient model of G.652 fiber (see blue curve in Fig. 9), a frequency dependent fiber loss profile is used as shown in Fig. 13(a) and we use 30 spans of 100 km. In addition, an optimized optical power profile is used at every span as shown in Fig. 13(b), assuming ideal optical power equalization. We consider C and L band Erbium doped fiber amplifiers at every span that ideally compensate for the span loss and with a uniform noise figure of 7 dB. The resulting accumulated ASE noise is loaded at the RX input. The optical launch power profile

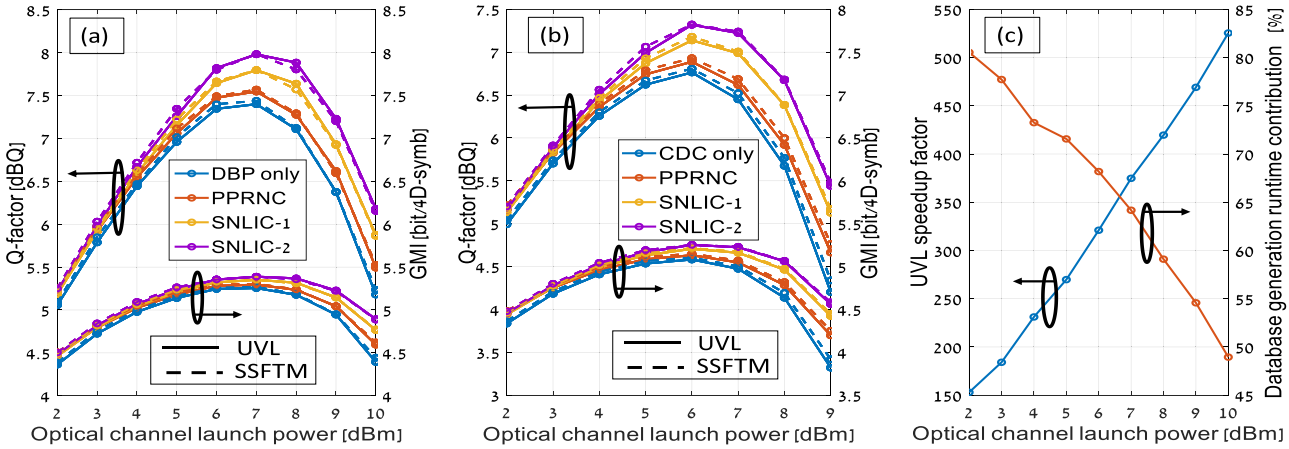

Fig. 11. Test scenario II: Q factor and GMI as a function of the optical channel launch power for UVL (solid lines) and SSFTM (dashed lines) with NOC=21with ideal DBP and SNLIC (a) and with CDC and SNLIC (b). UVL speedup factor (blue) and runtime contribution of the database generation in percentage (red) as a function of the optical launch power (c).

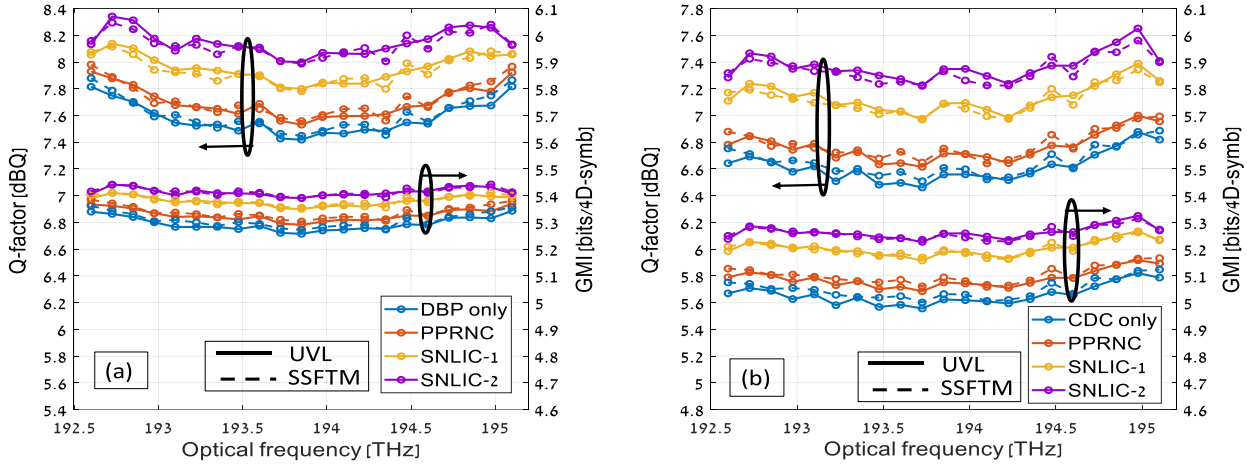

Fig. 12. Test scenario II: Q factor and GMI as a function of the COI optical frequency for UVL (solid lines) and SSFTM (dashed lines) with NOC = 21, optical channel launch power of 7 dBm and OSNR = 23 dB, NOC = 21 with ideal DBP and SNLIC (a) with CDC and SNLIC (b).

combined with ISRS, the intrinsic fiber loss profile and the ASE noise leads after 3000 km to the OSNR profile as shown in Fig. 13(c) with OSNR levels ranging from 21.2 dB to 21.55 dB over the whole 11 THz DWDM bandwidth. Fig. 13(d) shows the resulting XPM power and its first five ISI orders ($l = 0, +1, +2, +3, +4$) over the transmission bandwidth, after 3000 km and for COI received power normalized to 0 dBm. Due to the optical launch power profile, the XPM induced NLIN power increases towards the highest frequencies with a maximum reached at a COI frequency of 195.625 THz. The power contributions of the XPM induced NLIN orders decrease with the ISI order. The PPRN power contribution ranges from 12.2% to 17.6% of the total XPM power. Similarly, the powers of 1st and 2nd ISI orders of the NLIN contribute from 4.8% to 8.3% and from 2.7% to 5% of the total XPM induced NLIN power, respectively over the whole C+L band. The opposite sign ISI NLIN powers exhibit a similar contribution level.

Fig. 14 shows the real part of the ACFs of the *xx* and *xy* components of the ISI matrices $\mathbf{R}_l^{(n)}$ for $0 \leq l \leq 7$ at selected optical frequencies 186.5 THz, 191 THz and 196.25 THz[6]. These frequencies are indicated by dashed lines in Fig. 13(d) and correspond to optical channels located close to the two edges and the center of the ultra-wideband DWDM spectrum. The ACFs of the *yx* and *yy* components exhibit the same characteristics as their on-diagonal and off-diagonal component counterparts. The ACFs are obtained after SOS database generation and interpolation as explained in Section II-D. The relatively low PPRN power contribution described in Fig. 13(d) can be explained by the significant contribution of the higher ISI orders at the different optical frequencies, as shown in Fig. 14. Indeed, as described in the pulse collision model [36], transmission systems with lumped amplification and large span number increases the four-pulse collision occurrence since more incomplete collisions take place. Furthermore, at high baud-rates such as 120 GBd,

[6]According to Eq. (23) from Appendix A, the ACFs are obtained with IC decimation starting from IC frequency offset beyond 1.375 THz in order to ensure ACF accuracy within the symbol delay range of +/−1000 symbols.

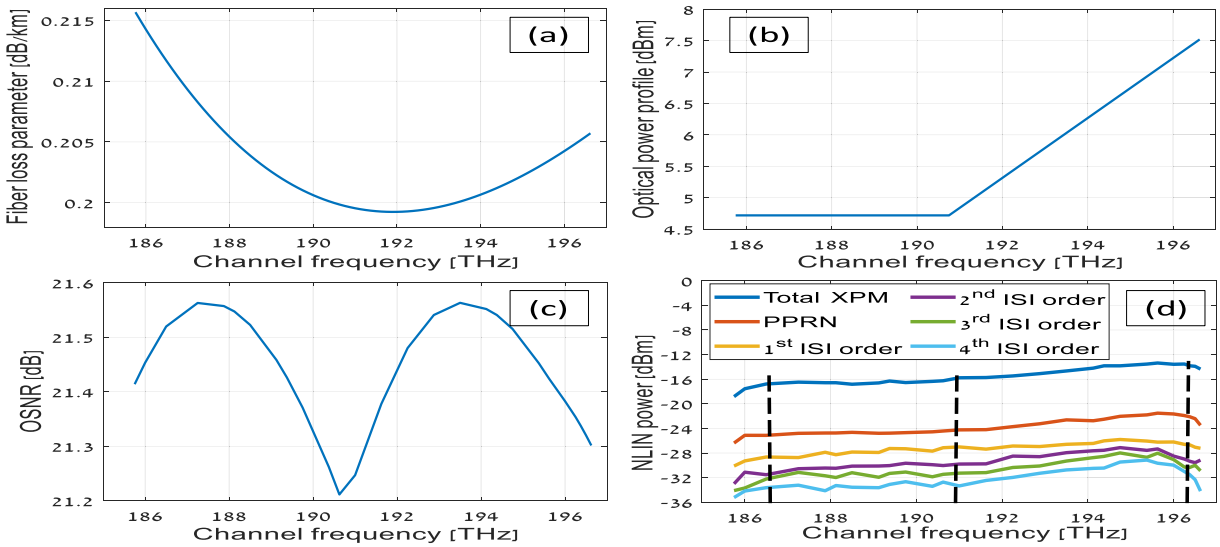

Fig. 13. Fiber loss parameter profile over C+L bands (a), Optimized optical launch power profile over C+L bands (b), OSNR profile after 30 × 100 km (c) and power profile of the total XPM induced NLIN and its five first ISI orders after 30 × 100 km normalized with respect to received COI power of 0 dBm (d). The dashed lines indicate the channel frequencies 186.5 THz, 191 THz and 196.25 THz.

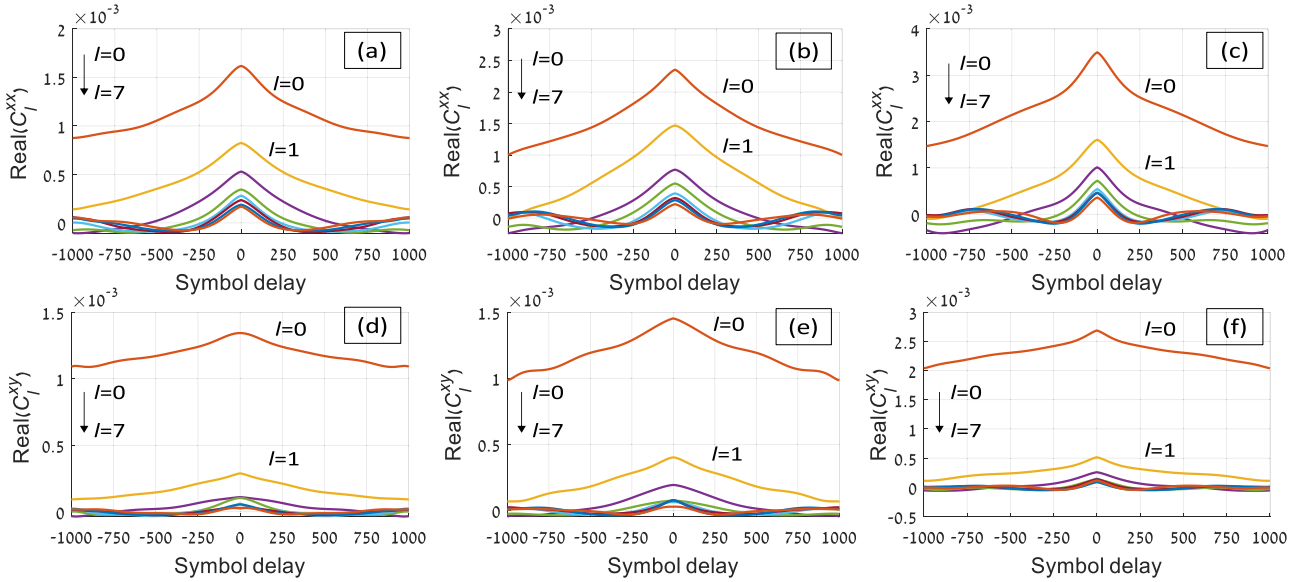

Fig. 14. Real parts of ACFs of the *xx* and *xy* components of the ISI matrices $\boldsymbol{R}_l^{(n)}$ for $0 \leq l \leq 7$ at optical frequencies 186.5 THz (a&d), 191 THz (b&e) and 196.25 THz (c&f).

the chromatic dispersion induced pulse broadening enhances incomplete collisions[7]. The correlation length is defined as the symbol delay leading to an amplitude reduction by 50% of the ACFs. Due to the large accumulated dispersion, the correlation length is of the order of multiple hundreds of symbols for ISI orders up to 2 with a symbol rate of 120 GBd. Especially for the PPRN, the correlation length exceeds 700 symbols for $\mathrm{R}_{0,xx}^{(n)}$ and is beyond 1000 symbol for $\mathrm{R}_{0,xy}^{(n)}$. The first order ISI exhibits a correlation length higher than 300 symbols and 450 symbols for $\mathrm{R}_{1,xx}^{(n)}$ and $\mathrm{R}_{1,xy}^{(n)}$ respectively. These extremely long correlation lengths should enable fairly effective the compensation for at least the $+/-1^{\mathrm{st}}$ order and to use large averaging windows in high BER regimes while tracking efficiently the changes in the ISI matrices.

Throughput performance assessments are provided in terms of GMI, used as Achievable Information Rate (AIR) [40]. The MIMO and CPR stages are identical to the ones used in the previous section while for nonlinear compensation schemes, we implement a filtered DBP (F-DBP) [41] with 1 step per span that

[7]The impact of the symbol rate on the XPM-induced NLIN, especially on the power contribution of the different ISI orders, requires a detailed analysis and will be presented in a future dedicated work.

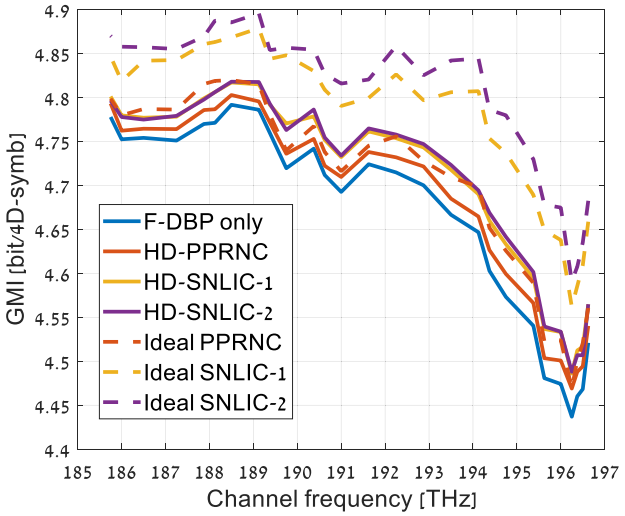

Fig. 15. GMI as a function of the COI optical frequency for C+L ultra-wideband system after different nonlinear compensation schemes. Solid lines refer to HD based nonlinear compensation schemes while dashed lines refer to ideal nonlinear compensation schemes.

TABLE II
TOTAL AIR IN C+L ULTRA WIDEBAND DWDM SYSTEM

Compensation scheme	Total AIR [Tb/s]	Total AIR increase with respect to F-DBP only
HD PPRNC	49.646	0.36%
HD SNLIC-1	49.891	0.86%
HD SNLIC-2	49.912	0.90%
Ideal PPRNC	49.852	0.78%
Ideal SNLIC-1	50.653	2.40%
Ideal SNLIC-2	50.919	2.93%

is used as non-ideal DBP approach and hard decision (HD) based PPRNC and SNLIC-1 and SNLIC-2 approaches, described in section III. By HD based nonlinear compensation, we mean that symbols after hard decision are used for NLIN estimation instead of using the transmitted symbols as in the genie mode approach applied in the previous section (in Eqs. (18-20)). Since with the UVL we have the ability to control the different XPM induced NLIN orders, we consider the following four cases:

Case 1: all the NLIN ISI orders are generated.
Case 2: all the NLIN ISI orders with $|l| \geq 1$ are generated.
Case 3: all the NLIN ISI orders with $|l| \geq 2$ are generated.
Case 4: all the NLIN ISI orders with $|l| \geq 3$ are generated.

While case 1 provides performance assessments for realistic link scenario including all the NLIN ISI orders, cases 2, 3 and 4 provide upper performance bounds for ideal PPRNC, ideal SNLIC-1 and ideal SNLIC-2 methods, respectively.

Fig. 15 shows the GMI of the 88 channels along the DWDM spectrum for the different cases. For all the cases, the GMI decreases towards the highest COI frequencies due to the combination of SPM and XPM noises and reaches a minimum at 196.25 THz. For case 1, a total AIR of 49.476 Tb/s is achieved when using F-DBP as the sole nonlinearity compensation scheme. As shown in Table II, after HD-PPRNC, the total AIR increases by merely 0.36% with respect to F-DBP while HD-SNLIC-1 and HD-SNLIC2 provide a moderate improvement of 0.86% and 0.9%, respectively. The very minor improvement provided by HD-PPRNC is due the fact that the PPRN contribution of the total XPM noise is very limited at high channel symbol rate. Even with ideal PPRNC, the increase in throughput is only 0.78%, growing to 2.4% and 2.93%, when ideal SNLIC-1 and SNLIC-2 are deployed, respectively. Ideal PPRNC and HD-PPRNC performances are very similar due to the long correlation length of the PPRN as shown in Fig. 14. In addition to HD-PPRNC which leads to a moderate gain, it is also possible to compensate the first ISI order of the XPM-induced NLIN due to its long correlation length. Additional compensation gain is negligible when applying HD-SNLIC-2. The decimation factor of the SOS database ranges from 66.6% to 71.3% according to the COI frequency.

Since computation of the full-bandwidth SSFTM is not feasible with a DWDM bandwidth of 11 THz, its runtime is roughly estimated by extrapolating the simulation runtime performed with 1/40$^{\text{th}}$ of the required symbol number. In a true SSFTM simulation, this number would need to be larger than the maximum walk-off between the edge channels of the DWDM spectrum [35]. Hence, for an 11 THz bandwidth DWDM system operating at 120 GBd, the sequence length should be 2^{19} symbols to avoid temporal aliasing and 217 samples per symbols should be used to avoid frequency aliasing. This leads to an estimated SUF of 35899.

V. CONCLUSION

The universal virtual lab has been proposed and demonstrated as a fast and accurate simulation tool for performance assessment in terms of Q-factor and GMI in nonlinear wideband DWDM systems, including the effect of ISRS. It has been shown to accurately account for practical transmitter and receiver imperfections and for spectral distortions following from the presence of WSS filters. As the UVL correctly models the temporal correlation properties of the NLIN, it accurately addresses scenarios including fast adaptive nonlinear equalization techniques. In all the tested cases, the error of the UVL is well contained within 0.1 dBQ and 0.09 bit/4D symbol for Q-factor and GMI, respectively. Huge speed-up factors relative to the full-bandwidth SSFTM follow both from the fundamental construction of the UVL method, as well as from the fact that unlike the SSFTM, it is suitable for using parallel processors. We demonstrate a speedup factor exceeding 1000 for the case of DWDM bandwidth of 3.375 THz (NOC = 45). When ISRS is negligible, power scaling of the generated database leads to speedup factors beyond 1000 for high optical launch power levels. When ISRS model needs to be included, the database must be generated for each launch power test case. We showed that in the case of a compressed SRS gain model with a DWDM bandwidth of 2.1 THz, a speed up factor of 375 is reached at the optimum optical power level. In the case of full C+L-band transmission, the speed advantage of the UVL over the SSFTM is estimated to be beyond 35000.

Using the UVL capabilities, we have also investigated the GMI and the total AIR in an ultra-wideband DWDM scenario with 88 $\times$ 120 GBd DP-PS-16QAM channels, accounting for

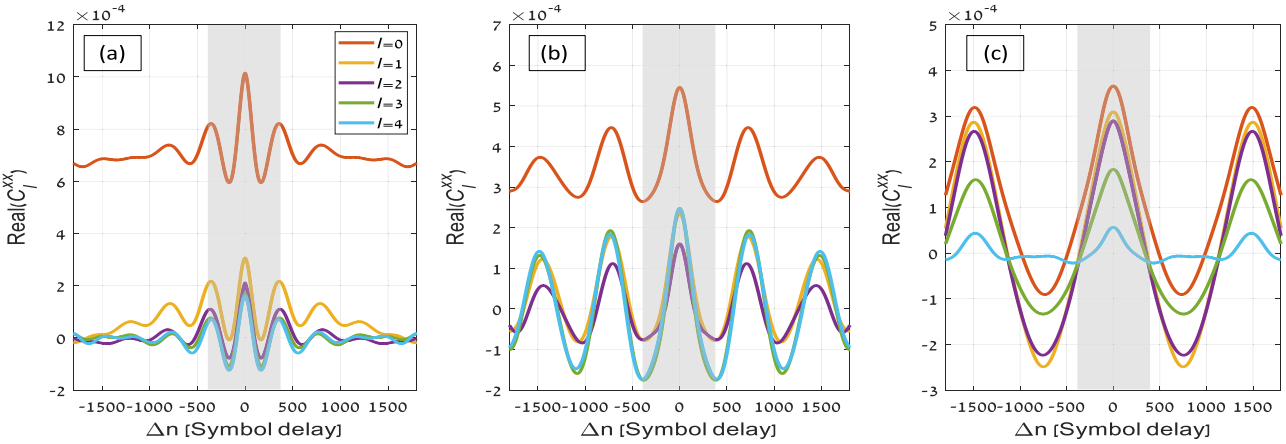

Fig. 16. Real part of the ACF of the $R_{l,xx}^{(n)}$ ISI component for $0 \leq l \leq 4$ with the 3000 km link scenario from section IV, with COI frequency of 196.25 THz and IC frequency offset of 250 GHz (a), 500 GHz (b) and 1 THz (c). The grey area refers to the symbol delay range used for the AR model generation with $\Delta \mathrm{n}_{\mathrm{MAX}} = 400$ symbols.

a bandwidth of 11 THz and an optimized optical launch power profile. The system includes transmission length of 30×100 km, non-ideal TX and RX architectures and distributed WSS as well as adaptive compensation schemes. We discussed the performances of practical nonlinear compensation schemes based on HD-PPRNC and HD-SNLIC approaches with respect to the compensation bounds at different NLIN ISI orders. Practical compensation strategies show that at 120 GBd, in addition to PPRNC which leads to a moderate gain, it is also possible to compensate the first ISI order of the XPM-induced NLIN due to its long correlation length, although the potential benefit from this compensation is seen to be fairly limited.

Appendix A

Frequency Offset Criterion for Database Decimation

With lumped optical amplification, oscillations are observed in the ACFs and PCFs. Indeed, when considering a single pulse of the IC as it propagates along the link, nonlinear interactions with the COI pulses occur at distances close to an amplifier and these nonlinear interactions become weaker, as the pulses propagate along the span because of the fiber attenuation. This means that the COI pulses nonlinearly collide with the IC pulse at similar distances from each amplifier in a comparable way and consequently, they are more correlated. These oscillations decay at long time symbol delays since this localization effect is reduced as the chromatic dispersion broadens the pulses very significantly. The use of distributed amplification reduces greatly these oscillations which vanish completely in an ideal lossless transmission system.

With lumped amplification, the IC oscillation period is expressed in symbol unit as

$$T_{IC} = 2\pi \left|\beta_2 \Delta f_{IC}\right| L_{\mathrm{span}} S_R, \tag{22}$$

where β_2 is the second order dispersion parameter at the COI frequency, $\Delta \mathrm{f}_{IC}$ is the frequency offset between the IC and the COI, S_R is the symbol rate and L_{span} is the span length that represents the lumped amplification period.

Fig. 16 shows the oscillation behavior of the ACF of the $\mathrm{R}_{l,xx}^{(n)}$ component for the 3000 km link scenario from Section IV with $0 \leq l \leq 4$ and with the COI frequency located at 196.25 THz. For IC frequency offsets of 250 GHz, 500 GHz and 1 THz, the oscillation period for $\mathrm{C}_0^{xx}(\Delta \mathrm{n})$ is 368, 736 and 1472 symbols, respectively in agreement with (22). The oscillation period obtained for higher ISI orders is almost equal to the case $l = 0$ but with very minor variations that result from more complex dynamics of the three and four-pulse collision scenarios involving two IC pulses [36]. A similar behavior is obtained for the ACF of the $\mathrm{R}_{l,xy}^{(n)}$ component. The grey area in Fig. 16 refers to the symbol delay range of interest, used for the AR model generation. To ensure that the SOS is accurate for a symbol delay range $|\Delta \mathrm{n}| \leq \Delta \mathrm{n}_{\mathrm{MAX}}$ after interpolation, we can choose the IC frequency offset beyond which the SOS evolves slowly within the symbol delay range of interest as a criterion for starting the decimation. This occurs when $\mathrm{T}_{IC}/2 \geq \Delta \mathrm{n}_{\mathrm{MAX}}$ and by using (22), we obtain the frequency offset criterion for decimation, $\Delta \mathrm{f}_{\mathrm{Decimation}}$ as follows

$$\left|\Delta f_{\mathrm{Decimation}}\right| \geq \frac{\Delta n_{\mathrm{MAX}}}{\pi \left|\beta_2\right| L_{\mathrm{span}} S_{\mathrm{R}}}. \tag{23}$$

For 120 GBd channels, spaced by 125 GHz and $\Delta \mathrm{n}_{\mathrm{MAX}} = 400$ symbols, (23) provides $\Delta \mathrm{f}_{\mathrm{Decimation}} = 500$ GHz.

References

[1] V. Lopez *et al.*, "Optimized design and challenges for C&L band optical line systems," *IEEE/OSA J. Lightw. Technol.*, vol. 38, no. 5, pp. 1080–1091, Mar. 2020.

[2] M. Cantono, R. Schmogrow, M. Newland, V. Vusirikala, and T. Hofmeister, "Opportunities and challenges of C+L transmission systems," *J. Lightw. Technol.*, vol. 38, no. 5, pp. 1050–1060, Mar. 2020.

[3] F. Hamaoka, "Ultra-wideband transmission and high-symbol rate signal handling technologies," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Paper W3E.1.

[4] G. Böcherer, P. Schulte, and F. Steiner, "Probabilistic shaping and forward error correction for fiber-optic communication systems," *IEEE/OSA J. Lightw. Technol.*, vol. 37, no. 2, pp. 230–244, Jan. 2019.

[5] Y. Yu *et al.*, "100.5 Tb/s MLC-CS-256QAM transmission over 600-km single mode fiber with C+L band EDFA," in *Proc. Asia Commun. Photon, Conf.*, 2018, pp. 1–3, doi: 10.1109/ACP.2018.8595747.

[6] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *IEEE/OSA J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[7] G. Agrawal, *Nonlinear Fiber Optics*. Amsterdam, The Netherlands: Elsevier, 2012.

[8] R. H. Stolen and E. P. Ippen, "Raman gain in glass optical waveguides," *Appl. Phys. Lett.*, vol. 22, pp. 276–278, 1973.

[9] M. Secondini and E. Forestieri, "Scope and limitations of the nonlinear Shannon limit," *IEEE/OSA J. Lightw. Technol.*, vol. 35, no. 4, pp. 893–902, Feb. 2017.

[10] E. Agrell, "Capacity bounds in optical communications," in *Proc. Eur. Conf. Opt. Commun.*, 2017, pp. 1–3, doi: 10.1109/ECOC.2017.8345871.

[11] D. Semrau, E. Sillekens, P. Bayvel, and R. I. Killey, "Modeling and mitigation of fiber nonlinearity in wideband optical signal transmission," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 12, no. 6, pp. C68–C76, Jun. 2020.

[12] O. Golani, M. Feder, and M. Shtaif, "NLIN mitigation using turbo equalization and an extended Kalman smoother," *IEEE/OSA J. Lightw. Technol.*, vol. 37, no. 9, pp. 1885–1892, May 2019.

[13] O. Golani, D. Pilori, F. P. P. Guiomar, G. Bosco, A. Carena, and M. Shtaif, "Correlated nonlinear phase-noise in multi-subcarrier systems: Modeling and mitigation," *IEEE/OSA J. Lightw. Technol.*, vol. 38, no. 6, pp. 1148–1156, Mar. 2020.

[14] C. Li *et al.*, "Inter-Channel fiber nonlinearity mitigation in high baud-rate optical communication systems," *IEEE/OSA J. Lightw. Technol.*, vol. 39, no. 6, pp. 1653–1661, Mar. 2021.

[15] D. Dahan and U. Mahlab, "In-band OSNR system margin and overall link impairment level monitoring techniques using an optical coherent receiver," *IEEE/OSA J. Lightw. Technol.*, vol. 33, no. 19, pp. 4063–4070, Oct. 2015.

[16] J. Auge, V. Curri, and E. L. Rouzic, "Open design for multi-vendor optical networks," in *Proc. Opt. Fiber Commun. Conf.*, 2019, Paper Th1I.2.

[17] D. Dahan and U. Mahlab, "Security threats and protection procedures for optical networks," *IET Optoelectron.*, vol. 11, no. 5, pp. 186–200, 2017.

[18] P. Serena, C. Lasagni, S. Musetti, and A. Bononi, "On numerical simulations of ultra-wideband long-haul optical communication systems," *IEEE/OSA J. Lightw. Technol.*, vol. 38, no. 5, pp. 1019–1031, Mar. 2019.

[19] D. Pilori, M. Cantono, A. Carena, and V. Curri, "FFSS: The fast fiber simulator software," in *Proc. 19th Int. Conf. Transp. Opt. Netw.*, 2017, PaperWe.B1.5.

[20] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri, "Modeling of the impact of non-linear propagation effects in uncompensated optical coherent transmission links," *IEEE/OSA J. Lightw. Technol.*, vol. 30, no. 10, pp. 1524–1539, May 2012.

[21] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Exp.*, vol. 22, no. 13, pp. 16335–16362, 2014.

[22] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-channel nonlinear interference noise in WDM systems: Modeling and mitigation," *IEEE/OSA J. Lightw. Technol.*, vol. 33, no. 5, pp. 1044–1053, Mar. 2015.

[23] M. Secondini and E. Forestieri, "Analytical fiber-optic channel model in the presence of cross-phase modulation," *IEEE Photon. Technol. Lett.*, vol. 24, no. 22, pp. 2016–2019, Nov. 2012.

[24] D. Semrau, E. Sillekens, R. I. Killey, and P. Bayvel, "A modulation format correction formula for the Gaussian noise model in the presence of inter-channel stimulated Raman scattering," *IEEE/OSA J. Lightw. Technol.*, vol. 37, no. 19, pp. 5122–5131, Oct. 2019.

[25] A. Bononi, R. Dar, M. Secondini, P. Serena, and P. Poggiolini, "Fiber nonlinearity and optical system performance," in *Springer Handbook of Optical Networks*. Berlin, Germany: Springer, 2020, pp. 287–351.

[26] P. Poggiolini *et al.*, "Non-linearity modeling for Gaussian-constellation systems at ultra-high symbol rates," in *Proc. Eur. Conf. Opt. Commun.*, 2018, Paper Tu4G.3.

[27] O. Golani, M. Feder, A. Mecozzi, and M. Shtaif, "Correlations and phase noise in NLIN-modelling and system implications," in *Proc. Opt. Fiber Commun. Conf.*, 2016, Paper W3I.2.

[28] O. Golani, M. Feder, and M. Shtaif, "Kalman-MLSE equalization for NLIN mitigation," *IEEE/OSA J. Lightw. Technol.*, vol. 36, no. 12, pp. 2541–2550, Jun. 2018.

[29] H. Rabbani *et al.*, "Analytical modeling of nonlinear fiber propagation for four dimensional symmetric constellations," *IEEE/OSA J. Lightw. Technol.*, vol. 39, no. 9, pp. 2704–2713, May 2021.

[30] O. Golani, R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Modeling the bit-error-rate performance of nonlinear fiber-optic systems," *IEEE/OSA J. Lightw. Technol.*, vol. 34, no. 15, pp. 3482–3489, Aug. 2016.

[31] D. Dahan, M. Zarubinsky, Y. Liang, O. Golani, and M. Shtaif, "Universal virtual lab: A fast and accurate simulation method for nonlinear DWDM system," in *Proc. Opt. Fiber Commun. Conf.*, 2021, Paper W6A.14.

[32] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Exp.*, vol. 22, pp. 16335–16362, 2014.

[33] J. N. Damask, *Polarization Optics in Telecommunications*. New York, NY, USA: Springer, 2005.

[34] O. Golani *et al.*, "Experimental characterization of nonlinear interference noise as a process of intersymbol interference," *Opt. Lett.*, vol. 43, no. 5, pp. 1123–1127, 2018.

[35] S. Musetti, P. Serena, and A. Bononi, "On the accuracy of split-step Fourier simulations for wideband nonlinear optical communications," *IEEE/OSA J. Lightw. Technol.*, vol. 36, no. 23, pp. 5669–5677, Dec. 2018.

[36] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications," *IEEE/OSA J. Lightw. Technol.*, vol. 34, no. 2, pp. 593–607, Jan. 2016.

[37] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats," *IEEE/OSA J. Lightw. Technol.*, vol. 35, no. 5, pp. 1125–1141, Mar. 2017.

[38] I. Fatadin, S. J. Savory, and D. Ives, "Compensation of quadrature imbalance in an optical QPSK coherent receiver," *IEEE Photon. Technol. Lett.*, vol. 20, no. 20, pp. 17333–11735, Oct. 2008.

[39] R. Dar and P. Winzer, "Nonlinear interference mitigation: Methods and potential gain," *IEEE/OSA J. Lightw. Technol.*, vol. 35, no. 4, pp. 903–930, Feb. 2017.

[40] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, "Achievable information rates for fiber optics: Applications and computations," *IEEE/OSA J. Lightw. Technol.*, vol. 36, no. 2, pp. 424–439, Jan. 2018.

[41] I. F. de Jauregui Ruiz, A. Ghazisaeidi, and G. Charlet, "Optimization rules and performance analysis of filtered digital backpropagation," in *Proc. Eur. Conf. Opt. Commun.*, 2015, pp. 1–3, doi: 10.1109/ECOC.2015.7342026.