Impact of filtering in optical coherent systems with finite-length equalizers: modeling and experimental verification

Impact of filtering in optical coherent systems with finite-length equalizers: modeling and experimental verification / Torres-Ferrera, P., Parisi, G., Rizzelli, G., Masanas, M., Gaudino, R., Fludger, C., Napoli, A.. - In: OPTICS EXPRESS. - ISSN 1094-4087. - 33:20(2025), pp. 41986-42004. [10.1364/oe.571379]

Research Article Vol. 33, No. 20 / 6 Oct 2025 / *Optics Express* 41986

Optics EXPRESS

Impact of filtering in optical coherent systems with finite-length equalizers: modeling and experimental verification

PABLO TORRES-FERRERA,[1,*] **GIUSEPPE PARISI,**[1] **GIUSEPPE RIZZELLI,**[2] **MIQUEL MASANAS,**[1] **ROBERTO GAUDINO,**[2] **CHRIS FLUDGER,**[3] **AND ANTONIO NAPOLI**[1]

[1] *Nokia NI, St.-Martin-Str. 76, 81541, Munich, Germany*
[2] *Politecnico di Torino, Turin, Italy*
[3] *Nokia NI, Nuremberg, Germany*
[*] *pablo.torresferrera@nokia.com*

Abstract: Modern ultra-high bit rate coherent transmission systems frequently operate under conditions where stringent optical and/or electrical filtering becomes a significant source of performance penalty. This is particularly evident when the end-to-end equivalent analog opto-electronic bandwidth falls below the transmitted signal symbol rate. Building upon our prior research, we introduce a novel semi-analytical model designed to quantify the performance degradation induced by filtering and colored-noise in coherent systems. Crucially, the model incorporates the effects of finite-length adaptive equalization at the receiver, thereby more accurately capturing the characteristics of realistic transceiver implementations. This developed tool offers not only a valuable and computationally efficient predictor of system performance, but also serves as a relevant resource for the design of networks and flexible transceivers. Furthermore, it facilitates rapid estimation of the impact of network variations, such as re-routing (which influences optical bandwidth) and power-saving mode operation (which affects idle equalizer taps), thus contributing to the development of real-time performance estimators. In this manuscript, we present, for the first time to the best of our knowledge, a comprehensive experimental verification of the proposed semi-analytical model. We consider the combined influence of Received Optical Power (ROP), Optical Signal-to-Noise-Ratio (OSNR), transceiver noise, and filtering on system performance. Additionally, we validate the model against extensive Monte Carlo numerical simulations for a distributed noise scenario. Our results demonstrate high model accuracy, achieving an estimation error below 0.15 dB compared to statistical numerical simulations, and, with respect to experimental measurements, an error below 0.35 dB in 90% of the cases, and below 0.6 dB in all the cases.

1. Introduction

The increasing adoption of coherent technology across all segments is becoming a key catalyst for an all-optical network integration. Coherent transponders were initially developed almost two decades ago for long-haul applications [1]. This technology simultaneously contributed to significantly simplifying optical system design (e.g., by removing dispersion compensation modules) and greatly increasing total throughput (e.g., by enabling high-order modulation formats). Next, coherent transceivers were progressively introduced to shorter links, such as metro-core networks and, more recently, inter-data center interconnections, targeting distances up to 80 km as proposed within the 400ZR implementation agreement. Recently, the ITU-T, in the context of the Very High Speed Passive Optical Network (PON) initiative (VHSP), is considering coherent technology as one of the possible options for next-generation 200G PON.

#571379 https://doi.org/10.1364/OE.571379
Journal © 2025 Received 20 Jun 2025; revised 15 Aug 2025; accepted 25 Aug 2025; published 24 Sep 2025

At the research level, numerous projects around the world have begun to envision a convergence between metro and access segments, which could be enabled by coherent transmission, and thus this option has become a focal point of several new research initiatives (e.g. [2–7]). The future envisioned metro access convergence, particularly in case the latter is deployed in a highly demanding physical layer, such as the PON architecture, introduces new opportunities and challenges that necessitate advanced modeling tools for effective link design. The first difference is that the ROP is typically much lower for access links compared to long-haul ones, due to the absence of in-line optical amplifiers in the access segment. In particular, for PON access networks, the ROP is usually well below -20 dBm. Second, in a metro + PON transmission system and for the downstream (DS) link, the received bit error ratio (BER) is thus primarily limited by a mix of available ROP and OSNR. Moreover, operation at low ROP may happen not only in optical access, but also on several new metro-regional architectures that are currently under investigation, particularly those proposing point-to-multipoint operation on horseshoe architectures using also optical splitters [8].

Another aspect that is becoming increasingly relevant for the design of the physical layer in future optical converged networks, particularly at higher symbol rates, is the growing impact of optical and electrical filtering. This is because the channels may traverse, for example, several Reconfigurable Optical Add Drop Multiplexers (ROADMs) and eventually be received (especially in DS scenarios toward "leaf" nodes) by a lower-cost coherent transceiver with limited opto-electronic bandwidth. An evolution of software modeling tools is then required to consider transmission conditions and links that are more sensitive to frequency-dependent impairments, such as those caused by the aforementioned filtering but also spectrally colored noise [9,10]. Currently, the most widely used quality of transmission (QoT) estimation tools [11] for network design, planning, dimensioning, and control, do not include the impact of filtering and colored noise inside the modulated signal bandwidth. However, several analytical models including these features have already been presented [12–20] and experimentally verified for both single-carrier [12,16,20] and multi-carrier [19] transmission. All of these models assume ideal adaptive feed-forward equalization (FFE) at the receiver, i.e, with an infinite number of taps, and most of them are based on the minimum mean square error equalizer (MMSEE) theory. This work introduces a significant enhancement to existing modeling tools by incorporating more realistic finite-tap adaptive equalizers, aligning more closely with those used in commercial coherent transceivers. A comparable approach is also discussed in [21] and verified against numerical simulations. In the present contribution, we expand the analytical model from [22] by introducing new considerations for the transceiver model based on empirical data, and detailing the formulation to model a multi-stage filtering system with distributed noise addition. We verified our proposed model against experimental measurements and numerical simulations.

The manuscript is organized as follows: In Sec. 2, we present a detailed review of previous articles on models for filtering penalty in coherent transmission systems. The experimental setup and the Monte Carlo numerical simulation environment used to assess the filtering penalty are then described in Sec. 3 and Sec. 4, respectively. The detailed description of the novel analytical model is presented in Sec. 5, whereas the comparison of the analytical results with experiments and numerical simulations are shown respectively in Sec. 6 and Sec. 7. Lastly, we draw some conclusions in Sec. 8.

2. Related literature, research gaps, and contribution

The mathematical foundations to introduce frequency-dependency to the well-known additive white gaussian noise (AWGN) model in a general digital communication system are presented in [23–27]. Semi-analytical formulations are derived, which allows using closed-form expressions to evaluate the system performance, directly estimating QoT metrics such as BER or Q-factor.

The framework developed is the so-called intersymbol interference (ISI)-channel model, as described in [25].

In the field of digital optical communications, initial analytical methods to estimate the impact of filtering on system performance were proposed in [28,29]. These methods rely on the calibration of the Signal-to-Noise-Ratio (SNR) versus OSNR relation, by approximating its nonlinear function in the presence of filtering through a Taylor series expansion. This requires fitting several Taylor coefficients, whose numbers increase with filter narrowing.

In [12], the ISI-channel model was introduced to estimate the impact of in-line filtering on a coherent optical communication system. The authors performed successful numerical and experimental verifications using a long-haul setup. The adaptive equalizer of the transceivers is modeled as a MMSEE with infinite taps. We name this approach "infinite MMSEE" throughout this article. In [25,30], it has been shown that adaptive fractional spaced equalizers (FSEs) that employs algorithms such as least mean square (LMS) and constant modulus algorithm (CMA), common in commercial transceivers, converge to the MMSEE solution. The work presented in [12] did not include the impact of frequency-dependent polarization impairments. In 2022, two independent works extended the approach proposed in [12], presenting an explicit analytical formulation that includes polarization effects [13,31]. The latter was expanded in [15], showing applications of the MMSEE-based model to both single-mode fiber long-haul systems and multi-mode fiber data-center interconnects. Later, an adaptation of the model based on MMSEE, addressing the optical metro access scenario, was presented in [16]. In that work, the low ROP regime of operation, typical in the PON segment, is taken into account, and the model is verified using experimental measurements. Subsequently, in [20], a model to estimate the impact of filtering, also on the metro access scenario, was presented. The authors follow a different approach to model the equalizer, assuming a zero-forcing (ZF) criterion. They show a consistent gap with the experimental measurements in the estimation of the Q-factor for OSNR values greater than 20 dB. This difference is ascribed to the fact that a ZF algorithm does not compromise ISI cancellation and noise enhancement, thus having in general a worse performance than a typical adaptive FSE using LMS or CMA for adaptation. In turn, in [18], the authors derive a filtering model for a similar scenario, but based on the infinite MMSEE approach as in [12,16]. The reported model is capable of following the trends obtained with numerical simulations in estimating the equalizer output SNR, in the presence of filtering, with a maximum error of 1 dB in its estimation.

Filtering impact models using the MMSEE approach have also been developed to cover Digital Subcarrier Multiplexing (DSCM) coherent systems, including the effect of quantization and clipping [14], and cross-talk among subcarriers [17]. A maximum error of 0.1 dB of the model in the estimation of the filtering penalty with respect to numerical simulations is shown in [17]. In [19], an experimental verification of the ISI-channel model considering infinite MMSEE applied to DSCM coherent systems is performed, achieving a maximum error of 0.5 dB in the estimation of the sensitivity penalty. The authors included in [19] a transceiver noise model following [28,32,33], where the impact of the ROP is taken into account, as well as a signal-dependent noise source that generates an intrinsic BER floor (i.e., an SNR ceiling), which is also used in the present contribution.

All of the previous approaches assume equalizers with infinite length. However, in [16,34] it is shown that an additional penalty associated with having a finite-tap equalizer is present under some filtering conditions. To analyze this phenomenon, a model that accounts for finite-tap equalization using MMSEE, for single-carrier coherent optical systems, is presented in [22]. We will refer to this approach as the "finite MMSEE" in the rest of this document. The authors verified the model versus numerical simulations, showing a maximum error of 0.1 dB in the estimation of the SNR at the equalizer output. A similar approach is followed in [21], to estimate the performance of Spatial Division Multiplexed (SDM) systems with finite-tap equalizers in the transceivers. In [21], a three-orders-of-magnitude reduction in computational cost is achieved with respect to numerical simulations using the semi-analytical model, with a maximum estimation error of ~0.3 dB.

The present contribution introduces a semi-analytical filtering model assuming finite-tap equalization at the receiver, aligning with real transceiver capabilities, and verifies it experimentally in single-carrier coherent optical systems. To the best of our knowledge, such verification has not been presented before.

In long-haul transmission, metro networks, and future metro-access, the optical signal traverses multiple stages of filtering and optical noise addition. Therefore, not only the signal is filtered, but also the noise. The optical noise added at each stage is filtered at the subsequent stages, arriving at the receiver as a colored-noise. This situation is referred to as distributed noise addition. Previous research has shown that the noise insertion point within a multi-stage system with optical filters plays a crucial role in determining the system performance [12,29]. The main findings of those works have been further confirmed in the simulations and experiments presented in [18,20,22]. In this contribution, we extend the model presented in [22], detailing the analytical formulation to address the distributed noise scenario in the multi-stage filtering plus noise system, having finite-length equalization in the transceivers, and we validate it against extensive Monte Carlo numerical simulations.

3. Experimental setup

We show in Fig. 1(a) the block diagram of the experimental testbed. We configured a transmission experiment influenced by optical filtering and optical noise, along with inherent electrical filtering and noise from the transceiver (TRX). We tested three real-time commercial coherent TRXs from different vendors: one operating at a net rate of 200 Gbps (the 200G TRX) and two at 400 Gbps (the 400G TRX-A and 400G TRX-B). The symbol rates (R_S) of the three TRXs are reported in Table 1. In all cases, the modulation format is dual-polarization (DP)-16-Quadrature Amplitude Modulation (QAM) and the pulse-shaping is carried out with a Root Raised Cosine (RRC) filter with the roll-off factor (ρ) reported in Table 1.

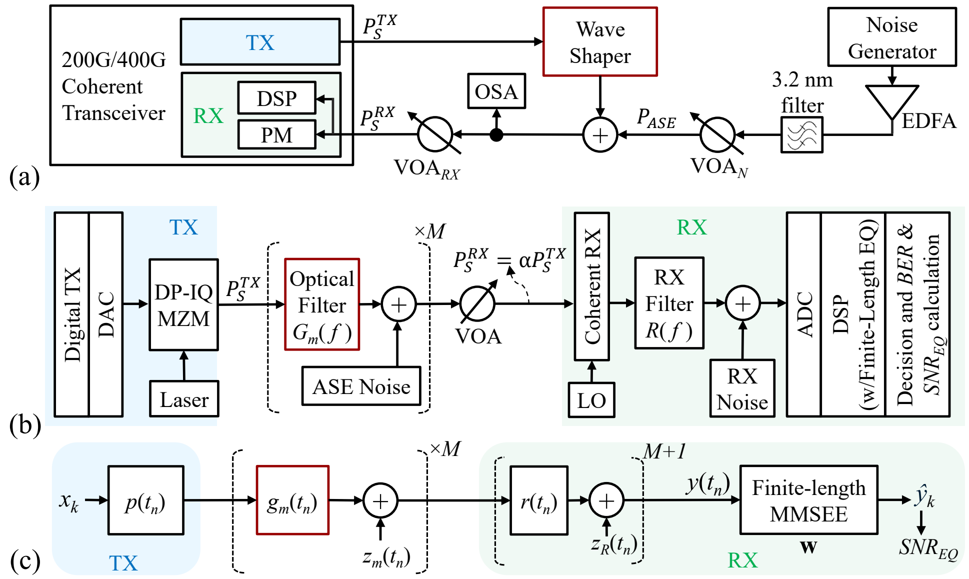

Fig. 1. (a) Experimental setup, (b) Simulation setup and (c) Model abstraction.

Table 1. Parameters of the tested transceivers from different vendors.

Parameter	200G TRX	400G TRX-A	400G TRX-B
Symbol rate, R_S	31.4 GBd	63 GBd	69.4 GBd
TX RRC roll-off, ρ	0.2	0.1	0.2
BER target, BER_T	1×10^{-2}	1×10^{-2}	2×10^{-2}
RX electrical filter ($R(f)$) bandwidth	37.6 GHz	68 GHz	67 GHz
RX electrical filter ($R(f)$) order	3	6	6
Number of taps of the equalizer tested in the model	4, 6, 8	14, 16, 18	14, 18, 28

We introduce optical filtering by loading different Super-Gaussian filter (SGF) power profiles [35], $T_{SGF}(f)$, into a programmable Wave Shaper. Each filtering profile is generated using Eq. (1), with nominal -3 dB bandwidth equal to B_{OF} and nominal order equal to η.

$$T_{SGF}(f)[\mathrm{dB}] = 20\log_{10}\left(G(f)\right) = 20\log_{10}\left(\exp\left(-\ln(\sqrt{2})\cdot(2|f|/B_{OF})^{2\eta}\right)\right) \tag{1}$$

We measured the actual filter profiles using an Optical Spectrum Analyzer (OSA), and these are presented in Fig. 2(a)-(c). From these measurements, we performed an analytical fit to the SGF profile defined in Eq. (1), to extract the actual parameters of the generated filters. The measured filter profiles and consequently their corresponding parameters deviate from the nominal values due to granularity limitations of the Wave Shaper in generating the optical filter. In Table 2 we report the relation between the parameters of the nominal and measured profiles, assuming SGFs in all cases. To ensure accurate comparisons between the measured and modeled system performance, presented in Section 6, we fed the analytical model with the measured values of B_{OF} and η, effectively incorporating the measured transfer functions into the model. It is important to note that our model is not limited to SGFs, but can be applied to any continuous and integrable filter shape, such as those typically used to model Wavelength Selective Switch (WSS) profiles [16,35–37]. Due to hardware availability, our experimental testbed contains a single stage of optical filtering and noise loading (i.e., $M = 1$), plus one stage of electrical filtering and noise addition intrinsic to the TRXs. The Amplified Spontaneous Emission (ASE) noise is added after optical filtering, which is the worst-case situation in terms of performance [12,28]. We dedicate Section 7 to extend our model verification considering multiple stages of optical filtering and distributed noise, by means of an extensive numerical simulation campaign.

Table 2. Filter parameters for nominal (N.) and corresponding measured (M.) profiles.

Used jointly w/	200G TRX			400G TRX-A								400G TRX-B					
B_{OF} [GHz] (N.)	25	29	32	50.4	53.5	56.7	63	56.7	63	56.7	63	54	57	60	63	70	72
η (N.)	2	2	2	3	3	3	3	6	6	9	9	3	3	3	3	3	3
B_{OF} [GHz] (M.)	22	26	29.5	49.2	52.4	55.7	62.2	54.4	61	53.9	60.6	52.7	55.8	58.8	61.8	69	71
η (M.)	1.3	1.5	1.66	2.4	2.5	2.5	2.6	3.4	3.6	3.8	4	2.4	2.5	2.6	2.6	2.6	2.7
B_{OF}/R_S (M.)	0.7	0.83	0.94	0.78	0.83	0.88	0.99	0.86	0.96	0.86	0.96	0.75	0.8	0.85	0.89	1	1.02

We control the power level of the ASE noise (P_{ASE}) by using a variable optical attenuator (VOA$_N$). Given that the average power of the transmitted signal P_S^{TX} remains constant, the OSNR is precisely determined by VOA$_N$. We measure the OSNR with an OSA, and report it on the basis of a reference bandwidth that matches the signal R_S. To properly compare performance with and without filtering, the reported OSNR, for a given value of P_{ASE}, is always the one measured when filtering is absent, as shown in Fig. 2(d). This is equivalent to defining the reported OSNR as: OSNR $= \alpha_{WS} \cdot P_S^{TX}/P_{ASE}$, with the P_{ASE} evaluated over a bandwidth equal to R_S. The term

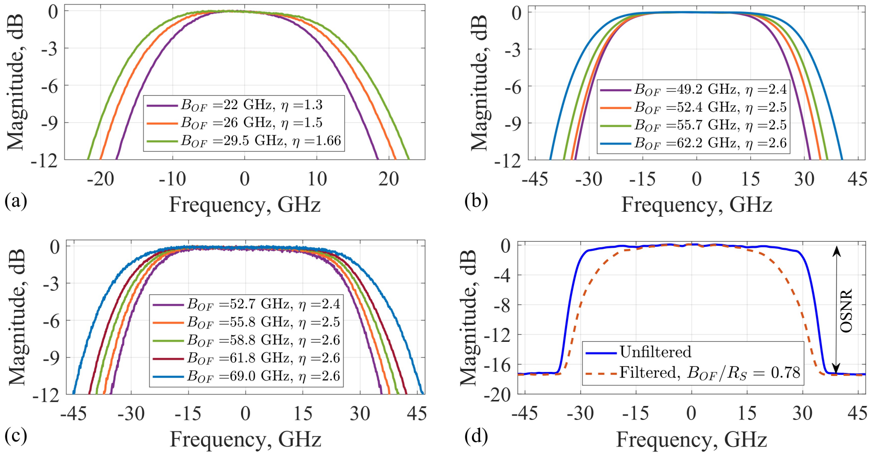

Fig. 2. (a)-(c) Samples of the measured filter profiles generated by the Wave Shaper, used in the: (a) 200G TRX experiment, (b) 400G TRX-A experiment, (c) 400G TRX-B experiment. (d) Received optical spectra (centered at f_s = 193.4 THz) without and with filtering (B_{OF} = 49.2 GHz) using the 400G TRX-A .

α_{WS} is the intrinsic insertion loss of the Wave Shaper, measured when loading an all-pass filter profile (i.e., $B_{OF} \rightarrow \infty$).

Finally, a second attenuator (VOA$_{RX}$) is used to set the signal average received optical power (P_S^{RX}). This quantity is evaluated as follows: $P_S^{RX} = \alpha_{RX} \cdot \alpha_{WS} \cdot P_S^{TX}$, where α_{RX} is the attenuation value set at VOA$_{RX}$. Note that P_S^{TX} and α_{WS} are constant values, irrespective of the optical filter shape, as stated before.

After coherent detection and Digital Signal Processing (DSP), the pre-FEC BER is calculated in real-time by the TRX. The commercial TRXs include a standard set of DSP algorithms [38] for signal recovery. We use the P_S^{RX} versus BER curves to evaluate the sensitivity at a given BER target (BER_T), under different filtering conditions. The baseline sensitivity, obtained in the case without optical filtering (i.e., $B_{OF} \rightarrow \infty$) and without ASE noise, is used as a zero-penalty reference to calculate the power penalties (ΔP, also termed sensitivity penalties in several works) when optical filtering and optical noise is introduced. Similarly, we use OSNR versus BER curves to evaluate the Required OSNR (ROSNR) at the given BER_T. A baseline ROSNR measured without optical filtering and with a high P_S^{RX} is used as a reference to calculate the ROSNR penalties.

4. Simulation setup

We show in Fig. 1(b) the setup used for Monte Carlo numerical simulations. We emulate a net 400 Gbps band-limited transmission using a coherent single-carrier TRX. The symbol rate is R_S = 64 GBd, the modulation format is DP-16-QAM and the pulse-shaping is a RRC with $\rho = 0.1$. After digital-to-analog conversion (DAC) and ideal electrical-to-optical conversion, the signal traverses $M-$stages of optical filtering plus the addition of optical noise prior to reception. Then, a $M + 1$ stage consisting of electrical filtering and electrical noise addition is considered before analog-to-digital conversion (ADC) and digital reception. We use SGFs of order η and -3 dB band-pass frequency equal to B_{OF} to study the degradation due to filtering. The frequency field response of each individual SGF, $G_m(f)$, is given by Eq. (2).

$$G_m(f) = \exp\left(-\ln(\sqrt{2}) \cdot (2|f|/B_{OF})^{2\eta}\right) \tag{2}$$

After each optical filter, the ASE noise is loaded. We properly set the OSNR and signal received optical power, P_S^{RX}, following the definitions indicated in the previous Section. Besides filtering and noise, other impairments, such as IQ imbalance, fiber nonlinearities, laser phase noise, and chromatic dispersion, are neglected in the simulation, being outside the scope of this work.

After ideal coherent detection, the received signal is resampled at $L = 2$ samples per symbol, emulating the ADC. Then, the digitized signal is processed through a standard set of DSP algorithms [38] for signal recovery. Because fiber propagation impairments and phase noise are neglected, the corresponding DSP algorithms, such as the fixed Chromatic Dispersion (CD) compensation, are turned off to save computational time. The main DSP algorithm for this work purposes is the adaptive FFE, which uses a training-aided decision-directed least mean square (DD-LMS) algorithm for adaptation. The adaptive equalizer is implemented as a fractional $T/2$-spaced equalizer in a finite impulse response (FIR) structure with a number of taps equal to $2 \times N_f$, where N_f is the length of the equalizer in symbol periods (T). A total of 2024 symbols are used to train the equalizer, and after training it is switched to a decision-directed mode, with a step size parameter $\mu = 10^{-4}$. Finally, after the decision, the BER is calculated by error counting and used to compute the corresponding SNR_{EQ}, using the following expression, valid for DP-16-QAM [39]:

$$SNR_{EQ} = 10 \cdot \text{erfc}^{-2}\left(\frac{8 \cdot BER}{3}\right) \tag{3}$$

5. Analytical model

We show in Fig. 1(c) the block diagram that we use in the analytical modeling. We abstract a linear optical coherent system, performing ideal optical-to-electrical conversion, and affected by $M-$stages of optical filtering plus ASE noise, followed by electrical filtering and the addition of electrical noise at the receiver. The model is defined in the time domain, based on the finite-length MMSEE theory presented in [23–25]. Due to the fact that a coherent receiver performs a linear map between the optical field and the electrical signals, it is not necessary to include opto-electrical conversions in the model. We did not include a matched filter in the receiver, since we assume a T/L fractional-spaced implementation of the MMSEE [25].

For simplicity, we start with the single-stage case, i. e., having $M = 1$. Therefore, we have $g(t_n) = g_1(t_n)$, $G(f_n) = G_1(f_n)$, and $z(t_n) = z_1(t_n)$. To simplify the analysis, we exclude the response of the receiver filter, $r(t_n)$, and the receiver noise source, $z_R(t_n)$. Later, we generalize for $M>1$ stages including all noise sources and filter impulse responses as reported in Fig. 1(c).

The receiver digitizes the output signal of the analog channel $y(t)$ taking L samples per symbol. The sampled channel output is the following:

$$y(t_n) = \sum_{i=-\infty}^{\infty} x_i \cdot h\left(t_n - iT\right) + z\left(t_n\right) \tag{4}$$

where $t_n = \frac{nT}{L} = \frac{(kL-j)T}{L}$, k is the symbol index, $j = 0, \ldots, L-1$ is the polyphase index [23,24], x_i is the i-th transmitted symbol, $h(t_n) = p(t_n) \circledast g(t_n)$ is the channel impulse response (here $\circledast$ is the convolution product), $p(t_n)$ is the pulse-shaping response, $g(t_n)$ is the optical filter response and $z(t_n)$ is the noise signal. We obtain $g(t_n)$ by applying the Inverse Fast Fourier Transform (IFFT) to $G(f_n)$, where $f_n = \frac{nL}{\nu T}$ and ν is the integer channel-memory. The sequences t_n and f_n account for the sampling range in the time and frequency simulation space, respectively.

Given N_f succesive L–tuples of samples of $y(t_n)$, and a discrete channel response $\mathbf{h} = [\mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_\nu]$ the oversampled vector representation of the channel at the equalizer input is

$\mathbf{Y}_k = \mathbf{H} \cdot \mathbf{X}_k + \mathbf{Z}_k$ [23,24], with:

$$\mathbf{X}_k \triangleq \left[x_k, x_{k-1}, \ldots, x_{k-N_f-\nu+1}\right]^T, \tag{5}$$

$$\mathbf{Y}_k \triangleq \left[\mathbf{y}_k, \mathbf{y}_{k-1}, \ldots, \mathbf{y}_{k-N_f+1}\right]^T, \tag{6}$$

$$\mathbf{Z}_k \triangleq \left[\mathbf{z}_k, \mathbf{z}_{k-1}, \ldots, \mathbf{z}_{k-N_f+1}\right]^T, \tag{7}$$

where $[\cdot]^T$ is the transpose operation while $\mathbf{X}_k$, $\mathbf{Y}_k$ and $\mathbf{Z}_k$ are the polyphase input/output signal and noise matrices, respectively, with vector rows:

$$\mathbf{y}_k = \begin{bmatrix} y(kT) \\ y(kT - \frac{T}{L}) \\ \vdots \\ y(kT - \frac{L-1}{L}T) \end{bmatrix}, \quad \mathbf{z}_k = \begin{bmatrix} z(kT) \\ z(kT - \frac{T}{L}) \\ \vdots \\ z(kT - \frac{L-1}{L}T) \end{bmatrix}, \quad \mathbf{h}_k = \begin{bmatrix} h(kT) \\ h(kT - \frac{T}{L}) \\ \vdots \\ h(kT - \frac{L-1}{L}T) \end{bmatrix}, \tag{8}$$

Finally, the channel autocorrelation-matrix is expressed as:

$$\mathbf{H} \triangleq \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \ldots & \mathbf{h}_\nu & 0 & \ldots & 0 \\ 0 & \mathbf{h}_0 & \mathbf{h}_1 & \ldots & \mathbf{h}_\nu & 0 & \ldots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \mathbf{h}_0 & \mathbf{h}_1 & \ldots & \mathbf{h}_\nu \end{bmatrix}, \tag{9}$$

with $\mathbf{H}$ being a $(N_f \cdot L) \times (N_f + \nu)$ Toeplitz matrix. The MMSEE processes each $\mathbf{Y}_k$ vector, to output the k-th equalized symbol $\hat{y}_k$ as: $\hat{y}_k = \langle \mathbf{w}^\dagger, \mathbf{Y}_k \rangle$, where $\mathbf{w}$ is the $(N_f L)$-dimensional vector containing the FIR MMSEE coefficients and $\langle \cdot, \cdot \rangle$ denotes the inner product. The equalizer output error, considering a properly chosen value of Δ for causal implementation, is then: $e_k = x_{k-\Delta} - \hat{y}_k$.

In this work, we neglect polarization-dependent impairments [15], thus assuming identical performance for X and Y polarizations. In the following, we show only the expressions for the X-polarization. The unbiased SNR at the output of the finite-length MMSEE is evaluated as:

$$SNR_{EQ} = \frac{\overline{\mathcal{E}}_x}{\sigma^2_{MMSEE}} - 1 \tag{10}$$

The term $\overline{\mathcal{E}}_x$ is the average energy of the signal, related to the average coherent power of the signal at the receiver input as follows: $P_S^{RX} = \overline{\mathcal{E}}_x \cdot R_S$. The term σ^2_{MMSEE} is the mean square error (MSE) of the equalizer output, evaluated as follows [23]:

$$\sigma^2_{MMSEE} = \mathbb{E}\{|e_k|^2\} = \overline{\mathcal{E}}_x - \mathbf{w} \cdot \mathbf{R}_{YY} \cdot \mathbf{w}^\dagger \tag{11}$$

where (†) is the transpose conjugation operation. For a sampling rate L/T, delay Δ, and length N_f symbol periods, $\mathbf{w}$ is evaluated as in Eq. (12) [24], when e_k and $\mathbf{Y}_k^*$ are on average orthogonal, which provides the minimum MSE.

$$\mathbf{w} = \mathbf{R}_{xY} \cdot \mathbf{R}_{YY}^{-1} \tag{12}$$

The terms $\mathbf{R}_{xY}$ and $\mathbf{R}_{YY}$ are the FIR MMSEE cross-correlation and autocorrelation matrix, respectively, calculated as follows [24]:

$$\mathbf{R}_{xY} = \mathbb{E}\{x_{k-\Delta} \cdot \mathbf{Y}_k^{\dagger}\} = \begin{bmatrix} 0 \cdots 0 & \overline{\mathcal{E}}_x & 0 \cdots 0 \end{bmatrix} \cdot \mathbf{H}^{\dagger} \tag{13}$$

$$\mathbf{R}_{YY} = \mathbb{E}\{\mathbf{Y}_k \cdot \mathbf{Y}_k^{\dagger}\} = \overline{\mathcal{E}}_x \cdot \mathbf{H} \cdot \mathbf{H}^{\dagger} + \mathbf{R}_{NN} \tag{14}$$

where $\mathbf{R}_{NN}$ is the fractionally spaced noise autocorrelation matrix and, when the noise is white, is $\mathbf{R}_{NN} = L \cdot N_0 \cdot I^{([N_f L],[N_f L])}/2$, with $I^{([N_f L],[N_f L])}$ being a $(N_f L) \times (N_f L)$ identity matrix and $N_0/2$ denoting the two-sided noise power spectral density (PSD).

For the general case of $M>1$, with distributed filtering and noise addition, we calculate the equivalent filter $g(t_n)$ and total colored-noise $z(t_n)$ from the sets of $g_m(t_n)$ and $z_m(t_n)$ including the action of $r(t_n)$ and $z_R(t_n)$ at the $M+1$ span, through the equivalent link transfer function [40] in the frequency domain:

$$\begin{aligned} Y(f_n) &= \left[\prod_{m=M+1}^{1} G_m(f_n)\right] P(f_n) X(f_n) + \left[\sum_{m=1}^{M+1}\left(\prod_{i=M+2}^{m+1} G_i(f_n)\right) Z_m(f_n)\right] \\ &= \tilde{H}(f_n) X(f_n) + \sum_{m=1}^{M+1} \tilde{G}_m(f_n) Z_m(f_n), \end{aligned} \tag{15}$$

with $G_{M+2}(f_n) = 1$, $G_{M+1}(f_n) = R(f_n)$ and $Z_{M+1}(f_n) = Z_R(f_n)$.

Equation (15) can be considered as the generalization of Eq. (4) to the case of $M+1$ filters in cascade expressed in frequency domain to simplify the analytical description. In Eq. (15), $X(f_n)$ and $Y(f_n)$ are the Fast Fourier Transform (FFT) of the input/output complex vector x_k and $y(t_n)$ in Fig. 1(c), $\tilde{H}(f_n) = \left[\prod_{m=M+1}^{1} G_m(f_n)\right] P(f_n)$ is the equivalent channel transfer function resulting from the product of all spans including the product of the filter transfer functions $P(f_n)$ and $R(f_n)$, while $Z_m(f_n)$ are the FFT of the noise sources $z_m(t_n)$ at the different stages. The terms $\tilde{G}_m = \prod_{M+2}^{i=m+1} G_i$ capture the cumulative filtering applied to the noise sources. This arises from the cascade of different filters (with impulse responses $g_m(t_n)$ and $r(t_n)$), which ultimately results in colored noise. We assume that x_k and z_m are independent and uncorrelated zero-mean Gaussian complex noise variables with variances $\sigma_x^2 = \overline{\mathcal{E}}_x$ and $\sigma_{z_m}^2$, respectively. The PSD of the receiver noise associated with $Z_R(f_n)$ is modeled as $S_R(f_n) = S_{R,0} + \beta \cdot P_R(f_n)/R_S$ [19,32,33]. In this equation, $S_{R,0}$ is the signal independent component of $S_R(f_n)$ (which is mainly accounted for thermal and shot noise) while $P_R(f_n)$ represents the received input power resolved by frequency and β is a parameter used to fit the model to experimental data. Consequently, $S_R(f_n)$ includes a noise component, $\beta \cdot P_R(f_n)/R_S$, which depends on the signal power as detailed in [32].

Finally, by converting $\tilde{H}(f_n)$ and $\tilde{G}_m(f_n)$ to the time domain via the IFFT, to obtain $\tilde{h}(t_n)$ and $\tilde{g}_m(t_n)$, we can determine the equivalent filter impulse responses affecting the signal and noise sources, respectively. These impulse responses then allow us to construct their respective autocorrelation matrices, $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{G}}_m$, which share a formal similarity with Eq. (9). We derived the generalization of Eq. (13) and Eq. (14) to the multi-stage scenario as:

$$\mathbf{R}_{xY} = \mathbb{E}\{x_{k-\Delta} \cdot \mathbf{Y}_k^{\dagger}\} = \begin{bmatrix} 0 \cdots 0 & \overline{\mathcal{E}}_x & 0 \cdots 0 \end{bmatrix} \cdot \tilde{\mathbf{H}}^{\dagger} \tag{16}$$

$$\mathbf{R}_{YY} = \mathbb{E}\{\mathbf{Y}_k \cdot \mathbf{Y}_k^{\dagger}\} = (1+\beta) \cdot \overline{\mathcal{E}}_x \cdot \tilde{\mathbf{H}} \cdot \tilde{\mathbf{H}}^{\dagger} + \sum_{m=1}^{M} \sigma_{z_m}^2 \cdot \tilde{\mathbf{G}}_m \cdot \tilde{\mathbf{G}}_m^{\dagger} + S_{R,0} \cdot I^{([N_f L],[N_f L])}, \tag{17}$$

with $\tilde{\mathbf{H}}$ being an $(N_f L) \times (N_f + \nu)$ Toeplitz matrix and $\tilde{\mathbf{G}}_m$ being an $(N_f L) \times (\nu L + N_f L)$ Toeplitz matrix. It is important to note that although x_k is a discrete-time sequence sampled at the symbol-rate, the noise sources are continuous-time signals that are sampled at L–times per symbol after undergoing convolution with the filters. For instance, for $L = 2$, the first column-vectors row of $\tilde{\mathbf{G}}_m$ is defined as in Eq. (18), where $\mathbf{0}$ is a $2 \times N_f$ zeros matrix.

$$\tilde{\mathbf{G}}_m(1:L=2,:) \triangleq \left[\tilde{\mathbf{g}}_{m,0,0}, \tilde{\mathbf{g}}_{m,0,1}, \tilde{\mathbf{g}}_{m,1,0}, \tilde{\mathbf{g}}_{m,1,1}, \ldots, \tilde{\mathbf{g}}_{m,\nu-1,0}, \tilde{\mathbf{g}}_{m,\nu-1,1}, \mathbf{0}\right]^T, \tag{18}$$

$$\text{with} \quad \tilde{\mathbf{g}}_{m,k,j} = \begin{bmatrix} \tilde{g}_m(kT + j\frac{T}{2}) \\ \tilde{g}_m(kT + j\frac{T}{2} - \frac{T}{2}) \end{bmatrix} \tag{19}$$

Using the generalized versions of $\mathbf{R}_{xY}$ and $\mathbf{R}_{YY}$, given in Eq. (16) and Eq. (17), respectively, we perform the procedure described before in Eqs. (10)–(12) to estimate SNR_{EQ}. Finally, the corresponding BER is evaluated using the following expression, valid for DP-16-QAM [39]:

$$BER = \frac{3}{8}\operatorname{erfc}\left(\sqrt{\frac{SNR_{EQ}}{10}}\right) \tag{20}$$

Before we conclude this section it is important to note a few key assumptions. Since our focus has been solely on filtering phenomena, we have assumed the absence of other linear impairments like Polarization Mode Dispersion (PMD), and Polarization Dependent Loss (PDL). We have also assumed that CD compensation is fully handled in a stage prior to the adaptive equalizer, which is the typical approach in modern coherent transceivers [38]. However, in cases of partial CD compensation, the inclusion of residual CD into the MMSEE model is straightforward as described in Appendix A.

6. Analytical model versus experimental measurements

6.1. Back-to-back characterization without optical filtering

To evaluate performance degradation due to filtering, we start the analysis by obtaining a baseline reference. As a first step, we characterized the performance of transceivers in the absence of optical filtering. In the following, this case is called the infinite-bandwidth case ($B_{OF} \to \infty$). We measured the pre-FEC BER as a function of both the ROP and the OSNR. From these measurements, we evaluated the sensitivity for each value of OSNR at the BER_T indicated in Table 1. For each transceiver, the best sensitivity value, obtained without ASE noise, is used as the reference for the evaluation of the power penalty due to filtering and OSNR degradation. In Fig. 3(a), Fig. 4(a) and Fig. 6(a), we show with markers the measured power penalty for the $B_{OF} \to \infty$ case, using the 200G TRX, 400G TRX-A and TRX-B, respectively.

In addition to obtaining a reference performance, we use the measurements without filtering to calibrate our analytical model. Through fitting, we estimate the β and $S_{R,0}$ parameters, associated with the signal dependent and signal independent noise terms of the TRX, respectively. We modeled the electrical filter at the receiver as an SGF with the parameters reported in Table 1.

We show in Fig. 3(a), Fig. 4(a) and Fig. 6(a) the estimated power penalty by the model for the $B_{OF} \to \infty$ case, using only lines. Three line styles–dotted, dashed, and dotted-dashed–correspond to different tap numbers, as detailed in Table 1. These values were chosen following our estimation of the range of taps for the operation of the tested TRXs. We also include a solid line, calculated using the infinite MMSEE assumption, following the model described in [19].

We observe that, even in the situation without optical filtering and high OSNR, there is a power penalty difference between the infinite MMSEE and the finite MMSEE model estimations, which is clearer in the 200G case. We can explain this difference because the intrinsic filtering, that is present at the transceiver, is not fully compensated with an equalizer with small number of taps.

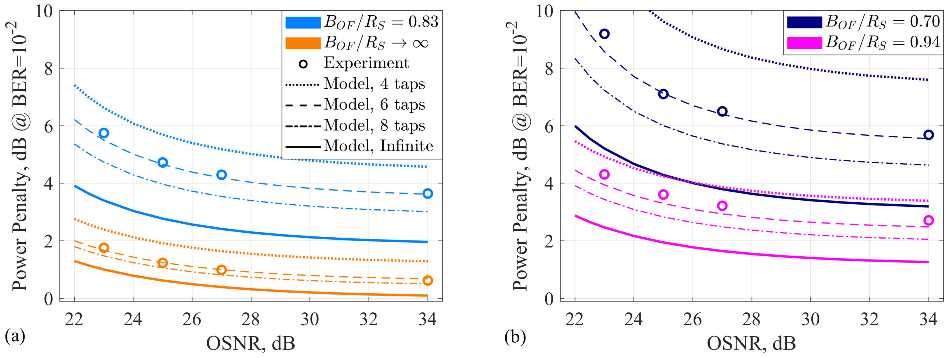

Fig. 3. Power penalty versus OSNR for several filter bandwidths and different number of taps, for the 200G transmission, with: (a) $B_{OF}/R_S = 0.83$ and ∞; (b) $B_{OF}/R_S = 0.7$ and 0.94. The line and marker code in legend of (a) applies to (b) as well.

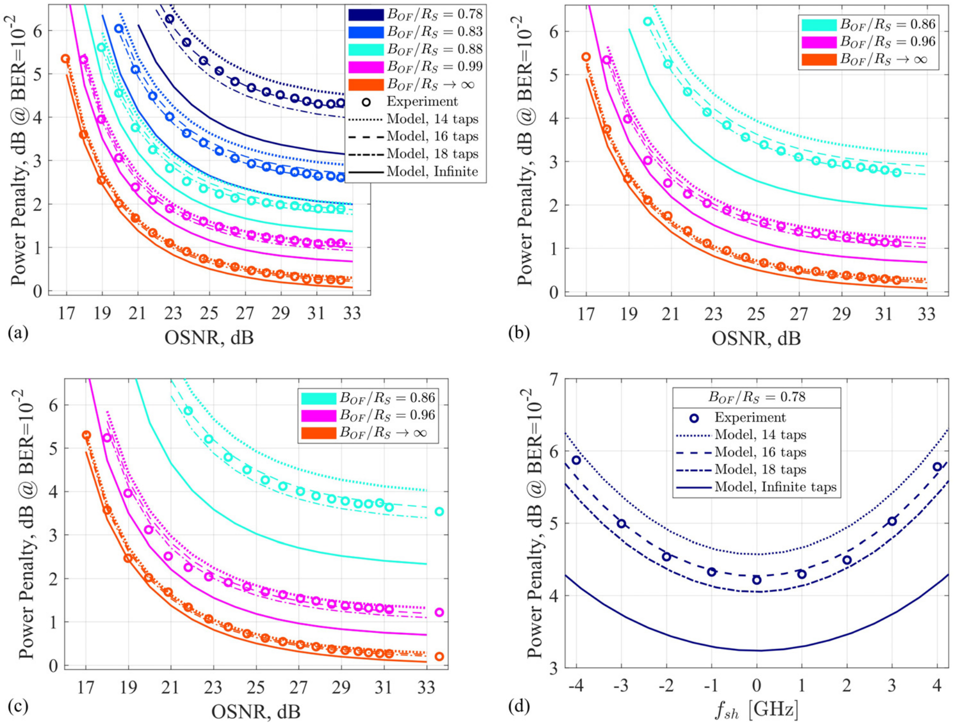

Fig. 4. (a)-(c) Power penalty versus OSNR for the 400G TRX-A transmission, for different number of taps and (a)-(c) for several measured filter bandwidths and nominal order equal to: (a) $\eta = 3$, (b) $\eta = 6$ and (c) $\eta = 9$; (d) and for a fixed filter with $B_{OF} = 49.2$ GHz and nominal $\eta = 3$, changing the frequency shift (f_{sh}) between the central frequencies of the signal and the filter. The line and marker code in legend of (a) applies to (b)-(c) as well.

It is worth pointing out for the reader that we are not including a dedicated matched filter at the receiver in the modeling, since the equalizer, in a fractional-spaced implementation, is meant to synthesize it, which is normally the case in the standard DSP of the commercial transceivers.

6.2. Performance penalty with optical filtering

In the following, we present the results when including optical filtering. For each of the measured filters presented in Table 2, we evaluated sensitivity versus OSNR relations, following the same procedure as in the previous subsection. Then, we calculate the corresponding power penalty ΔP with respect to the baseline sensitivity case without optical filtering and without optical noise. The results are displayed in Fig. 3, Fig. 4 and Fig. 6, using 200G TRX, 400G TRX-A and TRX-B, respectively. In all cases, we observe that the infinite MMSEE model is optimistic and cannot predict the penalty accurately. Conversely, the proposed finite MMSEE model, establishing the proper number of taps, is able to match the experimental measurements.

For the 200G experiment, we test three filters with a B_{OF} smaller than the symbol rate, including a stressed condition in which B_{OF} is only 70% of R_S. The results obtained, depicted in Fig. 3, show that, in all cases, the best match is obtained by setting 6 taps in the model, corresponding to an equalizer length of N_f =3 symbols. A sensitivity penalty of around 3 dB is measured when B_{OF} is 94% of R_S, for high OSNR values. When reducing the filter bandwidth to 70% of the symbol rate, this penalty increases to around 6 dB for high OSNRs. In this case, the infinite MMSEE model predicts a penalty that is 2 dB smaller, which exemplifies the need for a more accurate model. Regarding the finite MMSEE results for this scenario, if the number of taps is reduced to 4, the estimated penalty is approximately 2 dB higher. This indicates that, under some conditions, the performance is very sensitive to the equalizer length.

Regarding the 400G TRX-A experiment, we tested 18 filter profiles with B_{OF} ranging from 78% to 99% of R_S, and with nominal orders comprising η = 3, 6 and 9. As presented in Table 2, high-order filters produced by the Wave Shaper are observed to have an actual reduced order, which stems from the device granularity limitations. For simplicity, we refer to the nominal values of η in the following discussion. The results obtained are shown in Fig. 4(a), Fig. 4(b) and Fig. 4(c) for nominal η = 3, 6 and 9, respectively. From a comparison of Figs. 4(a)-(c), we observe that increasing the filter order causes a more pronounced performance degradation when using narrow filters (for instance: B_{OF}/R_S = 0.86 - 0.88) than when the filters have a bandwidth comparable to R_S (for instance B_{OF}/R_S = 0.96 - 0.99), or higher. In fact, considering a high OSNR, in the narrow-filter case the penalty increases from 2 to 3.5 dB, when increasing the nominal η from 3 to 9. In contrast, in the $B_{OF} \sim R_S$ case, ΔP remains around 1 dB. In Figs. 4(b)-(c), the curves corresponding to B_{OF}/R_S<0.86 values are not displayed because the BER_T is not achieved within the analyzed range of ROPs.

Consistent with the 200G experiment, the underestimation of ΔP by the infinite MMSEE model is clear, being equal to 1.5 dB for a filter with B_{OF}/R_S = 0.78 and nominal η = 3. Instead, the finite MMSEE model predicts well ΔP when configured with 16 taps. It should be noted that for most cases, particularly for filters with B_{OF}/R_S>0.85, the difference between the curves corresponding to 16 and 18 taps is minimal. Although this can lead to measured points appearing close to both curves, the finite MMSEE model with 16 taps consistently predicts ΔP within a maximum error range of 0.5 dB relative to the experimental measurements.

Figure 4(d) presents the power penalty as a function of the frequency shift (f_{sh}) between the central frequencies of the signal and the filter. This analysis considers a scenario with narrow filtering (B_{OF}/R_S= 0.78) and high OSNR (32 dB). Such an investigation is of practical interest, as real-world devices can experience frequency drifts due to temperature variations and device aging [9,10]. Therefore, accurate prediction of the resulting power penalty is essential. In fact, the filter generated for B_{OF}/R_S = 0.78 (shown in Fig. 2(b)), and used to obtain the results in Fig. 4(a), exhibits a f_{sh} =1 GHz relative to the signal. In Fig. 4(d) we tested f_{sh} values ranging

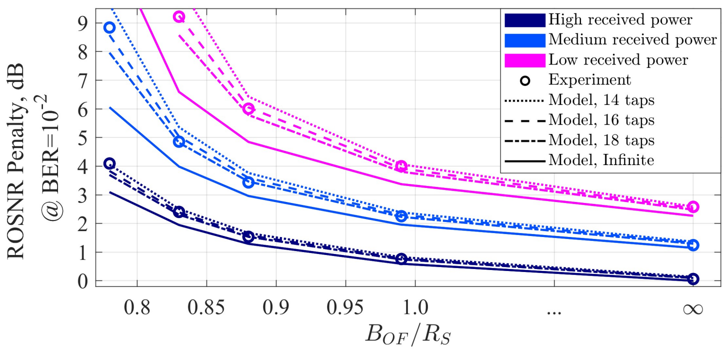

Fig. 5. Required OSNR penalty versus the optical filter -3 dB bandwidth to symbol rate ratio (B_{OF}/R_S), setting three values of received optical power (P_S^{RX}), for the 400G TRX-A transmission, using an optical filter with order η = 3.

from -4 GHz and 4 GHz. We observe a power penalty arising from this frequency shift, reaching 1.8 dB for $f_{sh} = \pm 4$ GHz, but remaining below 0.25 dB for $|f_{sh}| < 1.5$ GHz. The latter range of f_{sh} falls within the limits standardized by the ITU-T [9], and encompasses the statistics of measured values presented in [10]. In all instances, the finite MMSEE model accurately estimates the measured power penalty across the tested range of signal-to-filter frequency offsets.

Until now we have presented our results using the sensitivity penalty as a metric. This is a common metric useful to calculate the power budget in the optical distribution network of the access segment. We related the sensitivity penalty to the available OSNR to account for the constraints of the metro segment. To show the extended applicability of our model, in Fig. 5 we present our 400G TRX-A results using the ROSNR penalty as a metric. This metric is more commonly used in metro scenarios. We plot the ROSNR penalty as a function of the optical filter bandwidth normalized to the symbol rate. To generalize, we show the results for three values of the signal received optical power: a relatively high P_S^{RX} typical of in-line amplified metro links, a relatively low P_S^{RX} typical of access segments, and a medium value in between. We can observe that P_S^{RX} has a relevant impact on the performance, both with and without optical filtering. Decreasing P_S^{RX} reduces the corresponding electrical transceiver SNR, thus increasing the performance penalty. In the case without optical filtering (i.e. $B_{OF}/R_s \rightarrow \infty$), the observed ROSNR penalty is due only to this impairment. When filtering is introduced, the combination of transceiver SNR reduction and filtering penalty results in a ROSNR penalty that increases when reducing the received power. Our model accurately captures this behavior, showing good agreement across all filtering conditions and received power values with 16 taps. Conversely, the infinite MMSEE model fails to accurately predict measurements in these analyzed cases, particularly for low received power regimes typical of access and future metro-access networks.

Finally, Fig. 6 displays the results obtained from the 400G TRX-B experiment, using ΔP as a performance metric. We evaluated 6 filters with a nominal η = 3, and B_{OF} spanning from 75% to 102% of the symbol rate. The features of this TRX allows for a higher BER target of 2×10^{-2}, enabling the measurement of power penalty for a more stringently filtered case of B_{OF}/R_S = 0.75. We observe similar trends to those found with previous transceivers, which confirms the accuracy of the proposed finite MMSEE model, using 18 taps in this case. This 400G TRX-B, as reported in Table 1, operates with a higher symbol rate compared to 400G TRX-A, necessitating a couple more taps in its equalizer. For a better visualization, we present in Fig. 6 the results corresponding to 28 taps, as the performance difference between using 18, 20, or 24 taps is indiscernible in most of the cases.

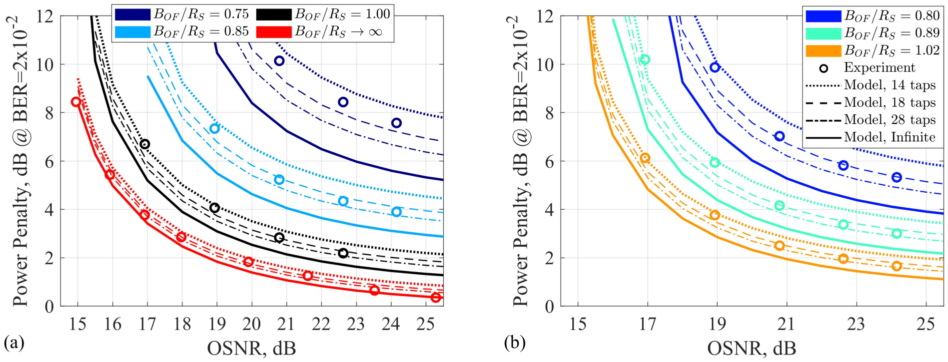

Fig. 6. Power penalty versus OSNR for several filter bandwidths and number of taps, for the 400G TRX-B transmission, with: (a) B_{OF}/R_S = 0.75, 0.85, 1 and ∞; (b) B_{OF}/R_S = 0.8, 0.89 and 1.02. Line and marker code in the legend of (b) applies to (a) as well.

Table 3 shows a summary of the power penalty estimation error (ΔP_{error}), when comparing the measurements versus the model predictions, for selected B_{OF}/R_S and OSNR values. In the model we set 6 taps, 16 taps and 18 taps, for the 200G, 400G-A and 400G-B cases, respectively. The ΔP_{error} is defined as the experimental power penalty minus the respective model prediction, in dB. We can observe a high model accuracy when using the finite MMSEE approach, with 90% of the ΔP_{error} values below 0.35 dB, and a maximum ΔP_{error} smaller than 0.6 dB. We attribute the estimation errors below 0.35 dB to the inherent measurement uncertainties of the instruments to meter ROP and OSNR. The largest estimation errors, higher than 0.35 dB, were observed under highly stressed conditions (narrow filtering and low OSNR) where the steep sensitivity penalty curve can amplify the effect of small measurement errors, leading to large variations in the sensitivity penalty. This region of operation, however, is of less practical interest due to the very large power penalties it produces.

Table 3. Comparison of Infinite (Inf.) and Finite (Fin.) MMSEE Models, in terms of power penalty estimation error (ΔP_{error}) compared to experimental measurements under varying B_{OF}/R_S ratios and nominal η = 3.

		B_{OF}/R_S Ratio							
		0.70 (200G)		0.83 (200G)		0.94 (200G)		∞	
		0.78 (400G-A)		0.83 (400G-A)		0.99 (400G-A)			
		0.75 (400G-B)		0.85 (400G-B)		1.02 (400G-B)			
	ΔP_{error} [dB]	**Inf.**	**Fin.**	**Inf.**	**Fin.**	**Inf.**	**Fin.**	**Inf.**	**Fin.**
200G	23 dB OSNR	3.91	0.54	2.34	0.25	1.80	0.35	0.76	0.08
	34 dB OSNR	2.48	0.13	1.68	0.03	1.44	0.23	0.52	-0.06
400G-A	22.7 dB OSNR	1.42	0.04	0.76	-0.09	0.3	-0.08	0.22	0.03
	32.3 dB OSNR	1.15	0.13	0.58	-0.12	0.4	0.08	0.14	-0.05
400G-B	20.78 dB OSNR	2.6	0.58	1.05	-0.1	0.48	-0.07	0.37	0.06
	24.15 dB OSNR	1.98	0.31	0.78	0.24	0.37	-0.15	0.06	-0.26

7. Analytical model versus numerical simulation results

In this section, we validate our finite MMSEE analytical model by comparing its predictions against Monte Carlo numerical simulations in various configurations, varying the filter parameters and distributed noise conditions. As a benchmark, we compare the predictions of the finite MMSEE model with the performance of an adaptive equalizer based on the DD-LMS algorithm. Under general assumptions of stationary and statistical independence between signal and noise sources, the two methods converge to a similar solution [25,30]. We remind the reader that the adaptive equalizer operates at a rate of 2 samples per symbol.

Figure 7 shows the SNR penalty as a function of the number of equalizer taps. To calculate the SNR penalty, the noise levels were set to define a reference baseline: the optical noise sources are set to achieve an output OSNR of 25 dB after the M-filter cascade (assuming $B_{OF} \to \infty$ in all filters), the SNR after the electrical RX filter is 25 dB, and the signal-dependent noise parameter is set to β = -20 dB. The selection of these values results in a reference baseline SNR of 17.9 dB.

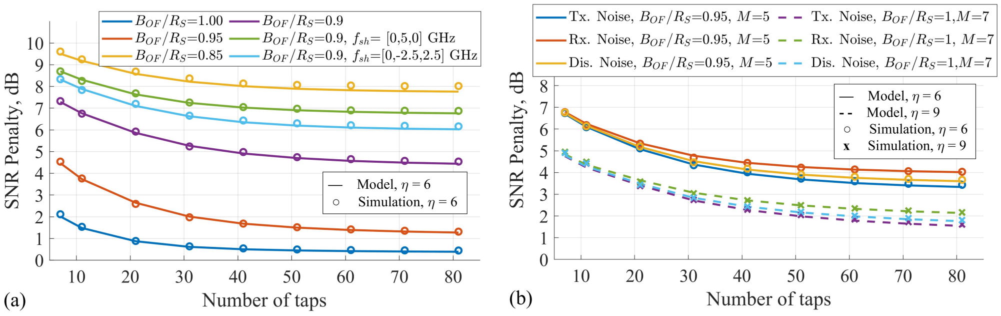

Fig. 7. SNR penalty versus taps for different filtering conditions. (a) M = 3, and equally distributed optical noise; (b) Without frequency shifts, fixed optical filter bandwidth and different conditions of optical noise addition

In Fig. 7(a), we test the model for a cascade of M = 3 filters with equally distributed optical noise. As expected, the results show that the SNR penalty is impacted by the bandwidth reduction. In particular, a penalty greater than 8 dB is predicted for filters with a bandwidth of 85% of the symbol rate. Furthermore, as discussed in the previous Section, a key strength of the model is its ability to predict penalties arising from frequency shifts between the filter and the signal carrier, as shown for two of the cases in Fig. 7(a). We use the following terminology to define a vector containing the frequency shifts of the filters in the system: $f_{sh} = [f_{sh,1}, f_{sh,2}, \ldots, f_{sh,m}, \ldots, f_{sh,M}]$, where $f_{sh,m}$ is the frequency offset of the m−th optical filter. In Fig. 7(b), the model is tested under more complex scenarios, varying the noise distribution and the number of filtering stages. The model's predictions confirm well-known results from the literature [12,18,22,29]: the worst-case penalty occurs when the noise is loaded at the end of the link (Rx. Noise), while the best case occurs when the noise is added at the beginning (Tx. Noise). Intermediate penalties are observed for distributed noise scenarios (Dis. Noise). Among them, the configuration of the noise equally distributed corresponds to the case shown in Fig. 7(b). In all the cases reported in Fig. 7, an excellent agreement is observed between the model and the simulations.

To further assess the model's accuracy, Fig. 8 presents a histogram of the SNR estimation error (SNR_{error}) for a cascade of M = 4 filters. For this test, the filter parameters were uniformly varied for each simulation run: the filter order η was chosen from 3 to 6, the frequency shift f_{sh} between -1.0 GHz and +1.0 GHz, and the filter -3 dB bandwidth B_{OF} between 90% and 100% of the symbol rate. A total of 500 iterations were run. The above choice of parameters reflect realistic scenarios reported in the literature [9,10]. The results show that the model is highly accurate, with an absolute estimation error consistently below 0.15 dB. The non-zero average SNR_{error} observed in the plots is an expected outcome, as the stochastic nature of the LMS algorithm causes it to approximate and oscillate around the optimal analytical solution of the MMSEE rather than converging precisely to it.

Fig. 9 provides an additional validation by comparing the internal equalizer characteristics for a cascade of $M = 4$ filters with $B = R_S$ and $\eta = 6$. Figure 9(a) shows the real part of the equalizer's impulse response ($\mathrm{Re}\{\mathbf{w}\}$), where the MMSEE model's prediction is well superimposed to the result from the LMS simulation. Similarly, Fig. 9(b) compares the estimated equivalent channel transfer function $\tilde{H}(f)$, from the tap coefficients $\mathbf{w}$. An excellent agreement is found among the MMSEE model, the LMS simulation, and the theoretical transfer function derived from Eq. (15).

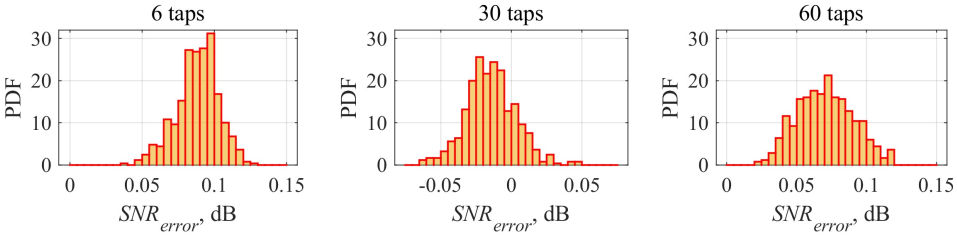

Fig. 8. Histogram of the SNR estimation error SNR_{error} for different number of taps

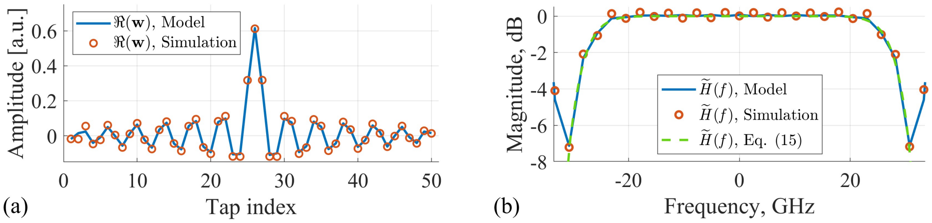

Fig. 9. (a) Impulse response of the equalizer (real-part ($\Re$)), (b) Estimation of the equivalent channel transfer function $\tilde{H}(f)$. In (a)-(b), the system parameters are: $M = 4$, $B_{OF}/R_S = 1$, $f_{sh} = [0,0,0,0]$ GHz and $\eta = 6$.

In summary, these results collectively demonstrate that the proposed analytical model accurately predicts the overall system penalty and reliably reproduces the physical behavior of both the channel and the equalizer.

Finally, we compute the average processing savings obtained using the semi-analytical model with respect to the numerical simulator based on error counting. This reduction in processing time is a key factor in favor of preferring a semi-analytical approach. Considering a channel memory of $\nu = 128$ in the model, and a PRBS of length of $2^{17} - 1$ symbols in the simulator (the same set of parameters used in Fig. 8), the model estimates a single SNR_{EQ} in average ~120 times faster than the numerical simulator. We refer the reader to Appendix B for a detailed complexity analysis comparing the Monte Carlo method with the proposed semi-analytical model.

8. Conclusions

We verified against experiments and numerical simulations a novel semi-analytical model to assess the performance penalty due to filtering and colored-noise in finite-length equalizer-based coherent systems. We tested the accuracy of the model for several filtering conditions. The model shows a high accuracy, achieving a maximum estimation error of 0.15 dB compared to extensive numerical simulations, and below 0.6 dB against experiments. We verify that the proposed model can reduce computational time by two-orders of magnitude compared to error-counting numerical simulations, building towards feasible real-time performance prediction.

It is important to note that the proposed model can incorporate the impact of other linear impairments. For instance, further extension of this work should address some open topics, like:

(i) The impact of residual CD that is not perfectly compensated by dedicated DSP algorithms prior to adaptive equalizer.

(ii) More generally, the impact of filters exhibiting non-constant group delay on the signal bandwidth. While this article employed filters with constant group delay, a more realistic non-constant group delay, although negligible with ideal infinitely-long equalizers, can introduce a performance penalty when using finite-length equalizers.

(iii) The impact of fiber propagation polarization impairments, such as PMD and PDL.

Appendix A: residual chromatic dispersion

In this Appendix we present an extension of our model including the effect of CD. Throughout this contribution we neglect the impact of fiber CD, by assuming that a frequency domain CD compensation (CDC) block, placed before the adaptive equalizer, fully compensates for the accumulated CD in the link [38]. However, a perfect CDC is not always achieved. Therefore, the effect of a residual amount of CD is still present in the signal that inputs the adaptive equalizer. In the following, we include this impairment into our finite MMSEE model. In Fig. 10 we show the model diagram considering this extension. Because CD acts as a global phase, it can be easily integrated into our model by updating $\tilde{H}(f_n)$ and $\tilde{G}_m(f_n)$ in Eq. (15) as follows:

$$\tilde{H}(f_n) \;\rightarrow\; \tilde{H}(f_n)e^{-i\frac{1}{2}(2\pi f_n)^2\left(\sum_{j=1}^{M}\beta_2\ell_j-2\beta_{CDC}\right)} \tag{21}$$

$$\tilde{G}_m(f_n) \;\rightarrow\; \tilde{G}_m(f_n)e^{-i\frac{1}{2}(2\pi f_n)^2\left(\sum_{j=m+1}^{M}\beta_2\ell_j-2\beta_{CDC}\right)} \tag{22}$$

Here, i is the imaginary unit, ℓ_j is the fiber length of the j-th span, β_2 is the fiber CD parameter and β_{CDC} represents the lumped CDC in ps^2 that simulates the effect of the static CDC pre-equalizer stage. Equation (21) describes the CD effect on the signal as it travels through the multi-stage link, while Eq. (22) accounts for the chromatic dispersion's impact on the different noise sources traversing varying lengths of optical fiber. We believe that a dedicated work devoted to characterize the impact of residual CD, when a finite-length adaptive equalizer is used, is an interesting open-topic for future extensions of the present contribution.

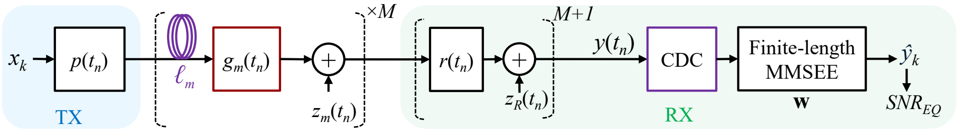

Fig. 10. Model abstraction to include chromatic dispersion effect.

Appendix B: complexity analysis

In this Appendix we discuss the complexity of the proposed analytical model and of the approach followed in the Monte Carlo numerical simulations using an LMS-based adaptive equalizer. We start with the analytical model. To compute the MSE ($\sigma^2_{\mathrm{MMSEE}}$) from Eq. (11), one must first sequentially compute the terms from Eq. (17), Eq. (16), and Eq. (12). The overall computational cost is the sum of these steps, but it is strongly dominated by the most complex operations.

The calculation of the cross-correlation vector R_{xY} from Eq. (16) is computationally inexpensive, with a linear complexity of $O(N_t)$ that can be neglected, where $N_t = N_f L$ is the adaptive equalizer number of taps. The dominant cost in this initial phase comes from building the autocorrelation matrix R_{YY} from Eq. (17), which requires matrix-matrix multiplications with a combined complexity on the order of $O(N_t^2 \cdot (N_f + \nu))$ for the $\tilde{\mathbf{H}} \cdot \tilde{\mathbf{H}}^\dagger$ products and M matrix multiplications, each with complexity $O(N_t^2 \cdot (N_t + \nu L))$ for the $\tilde{\mathbf{G}}_\mathbf{m} \cdot \tilde{\mathbf{G}}_\mathbf{m}^\dagger$ products. To find the coefficient vector $\mathbf{w}$, from Eq. (12), the primary cost is the inversion of the $N_t \times N_t$ matrix R_{YY}, which has a complexity of $O(N_t^3)$ with standard methods (i.e., Gaussian elimination). Other operations in this step are of a lower order and thus negligible. The final vector-matrix multiplication to get $\mathbf{w}$ is of a lower order, $O(N_t^2)$, and can thus be neglected. The calculation of σ^2_{MMSEE} itself, from Eq. (11), is dominated by a matrix-vector product of complexity $O(N_t^2)$, which is not dominant in the overall process. Comparing the steps, the construction of the matrix R_{YY} from Eq. (17) is the most computationally demanding operation, especially for multi-stage systems where $M \geq 1$. Its complexity grows faster than the matrix inversion cost. Therefore, the dominant cost for the entire process is determined by the matrix multiplications in Eq. (17). The approximate number of dominant operations is $\approx N_t^2 \left[N_f + \nu + M(N_t + \nu L)\right]$. The computational cost of the proposed semi-analytical model is a one-shot calculation, independent of the number of processed symbols. Its complexity is dominated by the construction and inversion of the system's correlation matrices, with an overall complexity on the order of $O(M \cdot N_t^3)$, where M is the number of system stages.

In contrast, the complexity of a Monte Carlo simulation using an iterative algorithm like LMS is linear with the number of processed symbols, $N_{PRBS} \approx 10^5$. The cost per symbol scales as $O(4N_t)$, considering four filters $\mathbf{w}_{xx}$, $\mathbf{w}_{xy}$, $\mathbf{w}_{yx}$, $\mathbf{w}_{yy}$ each of length N_t, resulting in a total complexity of $O(N_{PRBS} \cdot 4N_t)$. This fundamental difference in scaling with respect to the sequence length explains the significant computational advantage of the analytical approach, which avoids the time-consuming symbol-by-symbol loop inherent to simulations. Furthermore, we notice that the cost associated with the initial filtering is also higher in the simulation. Filtering the entire data sequence requires FFT-based convolution where the FFT length and number of operations depend on N_{PRBS} whereas in the analytical model it depends on the filter memory $\nu \ll N_{PRBS}$.

Funding. European Union Horizon Europe research and innovation program (GA No. 101092766 ALLEGRO).

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 gb/s coherent system," Opt. Express **16**(2), 873–879 (2008).
2. Y. Hu, A. Yan, J. Zhao, *et al.*, "All-optical metro-access integration network bidirectional transmission enabled by coherent digital subcarrier multiplexing," Journal of Optical Communications and Networking **17**(1), 58–70 (2025).
3. G. Rizzelli, M. Casasco, V. Ferrero, *et al.*, "Analysis and experimental demonstration of possible architectures for future coherent metro+PON converged networks [invited]," Journal of Optical Communications and Networking **17**(2), A142–A154 (2025).
4. F. Cavaliere, P. Iovanna, A. Bianchi, *et al.*, "Will a metro-access optical continuum ever fly? access network evolution trends and enabling technologies [invited]," Journal of Optical Communications and Networking **17**(2), A134–A141 (2025).
5. S. Al Zoubi and R. Gaudino, "Modeling and analysis of coherent metro + PON converged networks for ultra-high speed applications," Journal of Optical Communications and Networking **17**(4), 294–308 (2025).
6. S. Kaneko, Y. Kimura, K. Hara, *et al.*, "Metro/access converged all-photonics networks with remote control and management technologies accommodating various types of terminals [invited]," Journal of Optical Communications and Networking **17**(8), C105–C116 (2025).
7. G. Rizzelli, M. Casasco, A. Pagano, *et al.*, "Demonstration of converged Metro+Access bidirectional transmission using coherent transceivers and ROADMs," *2025 Optical Fiber Communications Conference and Exhibition (OFC)*, (2025), pp. 1–3.
8. M. M. Hosseini, J. Pedro, N. Costa, *et al.*, "Optimized design of horseshoe-and-spur filterless networks leveraging point-to-multipoint coherent pluggable transceivers," Journal of Optical Communications and Networking **16**(10), 969–980 (2024).
9. J. M. Fabrega, M. S. Moreolo, L. Martín, *et al.*, "On the filter narrowing issues in elastic optical networks," Journal of Optical Communications and Networking **8**(7), A23–A33 (2016).
10. J. Pan and S. Tibuleac, "Real-time ROADM filtering penalty characterization and generalized precompensation for flexible grid networks," IEEE Photonics J. **9**, 1–10 (2017).

11. P. Poggiolini and Y. Jiang, "Recent advances in the modeling of the impact of nonlinear fiber propagation effects on uncompensated coherent transmission systems," J. Lightwave Technol. **35**(3), 458–480 (2017).
12. I. F. de Jauregui Ruiz, A. Ghazisaeidi, T. Zami, *et al.*, "An accurate model for system performance analysis of optical fibre networks with in-line filtering," in *European Conference on Optical Communication (ECOC)*, (2019), pp. 1–4.
13. K. Benyahya, C. Simonneau, A. Ghazisaeldi, *et al.*, "Analysis of impact of polarization dependent loss in point to multi-point subsea communication systems," in *Optical Fiber Communications Conference and Exhibition (OFC)*, (2022), pp. 1–1.
14. A. Ghazisaeidi, K. Benyahya, and J. Renaudier, "On the potential benefits of entropy-loading for digital multicarrier signals: Analytical assessment of quantization noise and bandwidth limitation impairments," in *OptoElectr. Comm. Conf. (OECC)*, (2022), pp. 1–1.
15. G. Rizzelli, P. Torres-Ferrera, and R. Gaudino, "An analytical model for coherent transmission performance estimation after generic Jones matrices," J. Lightwave Technol. **41**(14), 4582–4589 (2023).
16. G. Rizzelli, P. Torres-Ferrera, A. Napoli, *et al.*, "Filtering effects characterization in metro-access networks and future high-speed coherent optical communication schemes," J. Lightwave Technol. **43**(3), 1114–1122 (2024).
17. P. Torres-Ferrera, G. Parisi, G. Rizzelli, *et al.*, "Analytical model for performance estimation of coherent digital subcarrier multiplexing systems affected by filtering, optical and electrical noise," in *24th International Conference on Transparent Optical Networks (ICTON)*, (2024), pp. 1–1.
18. E. Virgillito, S. Straullu, A. Castoldi, *et al.*, "Observation and modeling of filtering penalties in optical switched networks," in *24th International Conference on Transparent Optical Networks (ICTON)*, (2024), pp. 1–1.
19. P. Torres-Ferrera, G. Parisi, J. Sime, *et al.*, "Experimental verification of an analytical model of filtering impact on coherent digital subcarriers systems," in *Optical Fiber Communications Conference and Exhibition (OFC)*, (2025), pp. 1–1.
20. A. Rosso, E. Miotto, and E. Virgillito, other authors, "Encompassing filtering effects in transceiver models for converged metro-access networks," in *Optical Network Design and Modelling (ONDM)*, (2025), pp. 1–1.
21. R. Gholamipourfard, A. Ghazisaeidi, and R. S. B. Ospina, "A complexity-aware theoretical framework for the performance analysis of linear mimo equalizers in coherent spatial-division multiplexing transmission systems," J. Lightwave Technol. **43**(11), 5189–5201 (2025).
22. G. Parisi, P. Torres-Ferrera, G. Rizzelli, *et al.*, "Analytical model for filtering impact estimation of coherent systems with finite-length equalizer," in *Optical Network Design and Modelling (ONDM)*, (2025), pp. 1–1.
23. J. Cioffi, G. Dudevoir, M. Vedat Eyuboglu, *et al.*, "MMSE decision-feedback equalizers and coding. I. Equalization results," IEEE Transactions on Communications **43**(10), 2582–2594 (1995).
24. N. Al-Dhahir and J. Cioffi, "MMSE decision-feedback equalizers: finite-length results," IEEE Trans. Inf. Theory **41**(4), 961–975 (1995).
25. J. M. Cioffi, *Chapter 3 - Equalization*, in Standford Textbook on Digital Transmission, [Online] https://cioffi-group.stanford.edu/doc/book/chap3.pdf.
26. R. F. H. Fischer, *Linear Equalization*, Precoding and Signal Shaping for Digital Transmission (Wiley, 2002).
27. J. G. Proakis, *Digital Communications. Haup,3, Hauptbd. - 3. ed* (McGraw-Hill, 1995).
28. C. Delezoide, P. Ramantanis, and P. Layec, "On the performance prediction of optical transmission systems in presence of filtering," in *19th International Conference on Transparent Optical Networks (ICTON)*, (2017), pp. 1–4.
29. C. Delezoide, P. Ramantanis, and P. Layec, "Weighted filter penalty prediction for qot estimation," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, (2018), pp. 1–3.
30. H. Zeng and L. Tong, "Blind equalization using the constant modulus algorithm," **1**, 400–403 (1996).
31. G. Rizzelli, P. Torres-Ferrera, and R. Gaudino, "An analytical model for SNR prediction in coherent systems after generic Jones matrices," arXiv (2022).
32. L. Galdino, D. Semrau, D. Lavery, *et al.*, "On the limits of digital back-propagation in the presence of transceiver noise," Opt. Express **25**(4), 4564–4578 (2017).
33. G. Rizzelli Martella, A. Nespola, S. Straullu, *et al.*, "Scaling laws for unamplified coherent transmission in next-generation short-reach and access networks," J. Lightwave Technol. **39**(18), 5805–5814 (2021).
34. J. Pan and S. Tibuleac, "Filtering and crosstalk penalties for PDM-8QAM/16QAM super-channels in DWDM networks using broadcast-and-select and route-and-select ROADMs," in *2016 Optical Fiber Communications Conference and Exhibition (OFC)*, (2016), pp. 1–3.
35. T. A. Strasser and J. L. Wagener, "Wavelength-Selective Switches for ROADM applications," IEEE J. Sel. Top. Quantum Electron. **16**(5), 1150–1157 (2010).
36. F. Heismann, "System requirements for wss filter shape in cascaded roadm networks," in *2010 Conference on Optical Fiber Communication (OFC/NFOEC), collocated National Fiber Optic Engineers Conference*, (2010), pp. 1–3.
37. C. Pulikkaseril, L. A. Stewart, M. A. F. Roelens, *et al.*, "Spectral modeling of channel band shapes in wavelength selective switches," Opt. Express **19**(9), 8458–8470 (2011).
38. S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," IEEE J. Sel. Top. Quantum Electron. **16**(5), 1164–1179 (2010).
39. K. Cho and D. Yoon, "On the general ber expression of one- and two-dimensional amplitude modulations," IEEE Transactions on Communications **50**(7), 1074–1080 (2002).
40. A. Dumenil, E. Awwad, and C. Méasson, "PDL in optical links: A model analysis and a demonstration of a PDL-resilient modulation," Journal of Lightwave Tecnology **38**(18), 5017–5025 (2020).