This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

Impact of Chromatic Dispersion on Oversampled Digital Clock Recovery in Direct-Detection Systems: Analysis and Solutions

Patrick Matalla, *Student Member, IEEE,* Christian Koos, *ID* and Sebastian Randel, *Senior Member, IEEE*

***Abstract*—Cloud-based services such as the training of artificial intelligence (AI)/machine learning (ML) models in large-scale data centers as well as smart city applications such as the Internet of things (IoT), connected driverless vehicles, etc., are key drivers for ever higher data rates, especially in short-range fiber-optic links. Such optical transceivers employ intensity-modulation and direct-detection (IM/DD) to reduce the costs and power consumption and, therefore, chromatic dispersion (CD) is a nonlinear channel effect with respect to the received signal. With net data rates currently going up to 100 Gbit/s for passive optical networks (PONs) and 400 Gbit/s for Ethernet connections, CD is a main limitation in transmission performance at such high data rates. To reconstruct the signal impaired by CD, research has mainly discussed nonlinear equalizers and ML models, while, to the best of the authors knowledge, no research has been conducted on the influence of CD on non-data-aided, digital clock recovery in direct-detection systems. In this work, we investigate analytically and in simulation the effect of CD on oversampled digital clock recovery and find that for some dispersion values, clock recovery fails entirely. Based on our findings, we present two dispersion-tolerant clock recovery algorithms that work for non-return-to-zero (NRZ) and digital pulse-shaped signals as well as faster-than-Nyquist (FTN) signals. Finally, we validate our findings and algorithms in a 34-GBd four-level pulse amplitude modulation (PAM4) transmission experiment for NRZ and root-raised cosine (RRC) pulse-shaped signals and different accumulated dispersion values.**

***Index Terms*—Clock Recovery, Chromatic Dispersion, Direct-Detection, Faster-Than-Nyquist.**

I. Introduction

MODERN applications such as AI models trained and processed in the cloud, virtual reality, smart city applications, e.g., IoT, automated and connected driving, and many more, require dense and widespread communication networks that transmit high data rates with high reliability and low latency [1], [2], [3], [4]. Such applications are a key driver for ever higher data rates, especially over short distances less than few tens of kilometers. For such distances, optical fiber links based on IM/DD are the preferred choice due to their lower energy consumption and lower cost compared to coherent transceivers. The fields of applications of IM/DD transceivers range from PONs to datacenter interconnects (DCIs) and intra-datacenter (DC) optical links. All of these systems share a common challenge – as the data rate increases, CD becomes an increasingly limiting factor of the transmission performance, since the quadratic phase distortion of the optical field spectrum leads to a nonlinear channel effect when detecting the optical power.

In intra-DC and DCI, current research is considering 800 GE (Gigabit Ethernet) and 1.6 TE (Terabit Ethernet) as the next-generation standard, while discussion on 3.2 TE has already started [2], [3], [4]. Such high data rates are achieved by wavelength-division multiplexing (WDM), whereby it is more lucrative to achieve the highest possible data rate per wavelength, e.g., 16×200G or 8×400G for 3.2 TE. To provide an aggregated net data rate of 200 Gbit/s per wavelength up to fiber distances of 10 to 20 km in O-band, higher-order modulation formats, e.g., PAM4 up to eight-level pulse amplitude modulation (PAM8), are under consideration [5], [6], [7]. Despite the low dispersion in O-band, dispersion is a main limitation in transmission performance at such high symbol rates. The situation is similar for PONs, which are point-to-multipoint optical networks with passive optical splitters, used, for example, for broadband connectivity in cities [1] and 6G fronthaul [8] and leverage the technological advances in the Ethernet market. For the recently standardized 50G-PON, two-level pulse amplitude modulation (PAM2) modulation at 50 GBd is considered. With fiber lengths of up to 20 or even 40 km, CD is already a limiting factor here, which is why digital signal processing (DSP) equipped with adaptive equalizers is employed. The introduction of a DSP also allows conventional analog clock and data recovery (CDR) circuits to be migrated to the digital domain. For the 50G-PON standard, a dispersion range from −127 to 77 ps/nm is defined in O-band [9]. Current research is investigating the next potential IM/DD PON with a line rate of 100 to 125 Gbit/s using either 50 to 62.5 GBd PAM4 or 100 to 125 GBd PAM2. Depending on the choice of wavelength, a dispersion tolerance from −120 to 80 ps/nm is required [10].

Previous research has mainly focused on system aspects and feasibility studies for IM/DD systems at such high data rates. A major focus here was the signal attenuation in the spectral dips caused by CD in direct-detection links and on counteracting this with sophisticated nonlinear equalizers and ML methods [11], [12]. The effect of CD on the digital clock recovery is investigated in several works for coherent transmission [13], [14]. However, to the best of the authors'

This work was supported by the German Federal Ministry of Education and Research (in German: Bundesministerium für Bildung und Forschung) in the project HYPERCORE under Grant 16KIS2103 and project PONTROSA under Grant 16KIS2092K.

P. Matalla, C. Koos, and S. Randel are with the Institute of Photonics and Quantum Electronics, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany. (e-mail: patrick.matalla@kit.edu; christian.koos@kit.edu; sebastian.randel@kit.edu).

Manuscript received April 19, 2005; revised August 26, 2015.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

knowledge, no research has been done on the effect of CD on the clock recovery in direct-detection systems.

This work is structured in two parts. In the first part, we analyze and simulate the effect of CD on oversampled digital clock recovery in direct-detection systems. We holistically investigate the effect for signals with NRZ pulses as well as signals with RRC pulse shaping. Building on the findings from the first part, we then propose two CD-tolerant clock recovery algorithms that work for NRZ or high roll-off signals as well as for low roll-off or extremely bandwidth-limited signals such as FTN signals. Finally, we confirm our analysis and the viability of the proposed clock recovery algorithms in a 34-GBd PAM4 transmission experiment over accumulated dispersion values of up to 600 ps/nm.

II. Impact of Chromatic Dispersion on Digital Clock Recovery

A. Signal Propagation through Chromatic Dispersive Channel

To understand the effect of CD on the clock recovery for direct reception at the receiver end, we first consider the influence of CD on the received intensity-modulated signal. In [15], [16], the received optical power impaired by CD is analytically derived. We consider a zero-mean, real-valued transmit signal $s(t)$ in time domain, which is normalized to the mean optical power P_0 and $\min\{s(t)\} \geq -1$. The transmitted optical power can then be described as $P_{\text{tx}}(t) = P_0(1+s(t))$. Considering chirp-free modulation and neglecting any noise, we can assume a constant optical phase and express the transmitted complex amplitude with respect to optical center frequency as the square-root of the optical power, i.e.,

$$E_{\text{tx}}(t) = \sqrt{P_0\,(1+s(t))} \tag{1}$$

with the extinction ratio (ER) in decibels given as

$$\text{ER}_{\text{dB}} = 10\log_{10}\left(\frac{1+\max(s(t))}{1+\min(s(t))}\right). \tag{2}$$

The square-root can be expressed as a power series resulting in

$$E_{\text{tx}}(t) = \sqrt{P_0}\left(1+\sum_{n=1}^{\infty}\alpha_n s^n(t)\right), \tag{3}$$

with coefficients

$$\alpha_n = \frac{(-1)^{n-1}(2n)!}{4^n(n!)^2(2n-1)} \qquad \text{for} \quad n\in\mathbb{N}^+. \tag{4}$$

As the optical signal propagates through a dispersive medium, it experiences CD. The optical channel $\tilde{\underline{h}}_{\text{CD}}(f)$ can be described by an allpass filter with quadratic spectral phase in frequency-domain (FD) as

$$\tilde{\underline{h}}_{\text{CD}}(f) = \exp\left(\mathrm{j}\,\frac{\pi}{c}\lambda^2 f^2 L D_{\text{CD}}\right) = \exp\left(\mathrm{j}\,\zeta(f)\right), \tag{5}$$

where D_{CD} is the CD coefficient, f the equivalent baseband frequency of the signal, λ the center wavelength, L the fiber length, c the speed of light, the tilde indicates being in the Fourier domain, and the underline denotes a complex value/signal. For notational convenience, we abbreviate the argument by $\zeta(f) = \pi\lambda^2 f^2 L D_{\text{CD}}/c$. Since the real and imaginary parts of the channel frequency response are both even real functions, the real and imaginary parts of the impulse response $\underline{h}_{\text{CD}}(t) = \mathcal{F}^{-1}\left\{\tilde{\underline{h}}_{\text{CD}}(f)\right\}$ can be expressed as

$$\begin{aligned}\Re\{\underline{h}_{\text{CD}}(t)\} &= \mathcal{F}^{-1}\left\{\Re\left\{\tilde{\underline{h}}_{\text{CD}}(f)\right\}\right\} = \mathcal{F}^{-1}\left\{\cos\left(\zeta(f)\right)\right\}\\ \Im\{\underline{h}_{\text{CD}}(t)\} &= \mathcal{F}^{-1}\left\{\Im\left\{\tilde{\underline{h}}_{\text{CD}}(f)\right\}\right\} = \mathcal{F}^{-1}\left\{\sin\left(\zeta(f)\right)\right\}.\end{aligned} \tag{6}$$

The received optical power then results as the absolute square of the received optical field as

$$P_{\text{rx}}(t) = \left(E_{\text{tx}}(t) * \Re\{\underline{h}_{\text{CD}}(t)\}\right)^2 + \left(E_{\text{tx}}(t) * \Im\{\underline{h}_{\text{CD}}(t)\}\right)^2, \tag{7}$$

where $*$ denotes the convolution operation. Since the Fourier transform of the constant term corresponds to the Dirac impulse $\delta(f)$ and using eq. (6), the convolution of the direct-current term with the real and imaginary part of the CD impulse response results to one and zero, respectively, i.e.,

$$\begin{aligned}\mathcal{F}\{1 * \Re\{\underline{h}_{\text{CD}}(t)\}\} &= \delta(f)\cos\left(\zeta(f)\right) = 1\\ \mathcal{F}\{1 * \Im\{\underline{h}_{\text{CD}}(t)\}\} &= \delta(f)\sin\left(\zeta(f)\right) = 0.\end{aligned} \tag{8}$$

The received electrical signal $x(t) \propto P_{\text{rx}}(t)$ results in [16]

$$x(t) \propto 1 + s(t) * \Re\{\underline{h}_{\text{CD}}(t)\} + \varepsilon(t) \tag{9}$$

with

$$\begin{aligned}\varepsilon(t) = {} & 2\sum_{n=2}^{\infty}\alpha_n s^n(t) * \Re\{\underline{h}_{\text{CD}}(t)\}\\ & + \left[\left(\sum_{n=1}^{\infty}\alpha_n s^n(t) * \Re\{\underline{h}_{\text{CD}}(t)\}\right)^2\right.\\ & \left. + \left(\sum_{m=1}^{\infty}\alpha_m s^m(t) * \Im\{\underline{h}_{\text{CD}}(t)\}\right)^2\right].\end{aligned} \tag{10}$$

The first term in eq. (9) is a constant direct-current offset, while the second term is referred to as power fading (PF). Considering the Fourier transform of the received signal $\tilde{\underline{x}}(f)$

$$\tilde{\underline{x}}(f) \propto \delta(f) + \tilde{\underline{s}}(f)\tilde{h}_{\text{PF}}(f) + \tilde{\underline{\varepsilon}}(f) \tag{11}$$

with the frequency response

$$\tilde{h}_{\text{PF}}(f) = \Re\left\{\tilde{\underline{h}}_{\text{CD}}(f)\right\} = \cos\left(\zeta(f)\right), \tag{12}$$

we find that the power fading effect shapes the spectrum of the received signal and imposes nulls at frequencies f_m

$$f_m = \pm\sqrt{\frac{(1+2m)c}{2\lambda^2 L D_{\text{CD}}}} \qquad \text{for} \quad m\in\mathbb{N}^0. \tag{13}$$

Note, that the impact of the power fading depends on the spectral extent of the signal $s(t)$ and the accumulated dispersion LD_{CD}. $\varepsilon(t)$ comprises the nonlinear interference terms, which include the power fading of the signal's power series and signal-signal beat interferences.

The simulation setup used to investigate the effect of CD is shown in Fig. 1. Here, we simulate a 112-GBd PAM4 signal with NRZ pulse shape at 8-fold oversampling. Afterwards, we apply a 5th-order Bessel lowpass filter with a 3-dB bandwidth

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

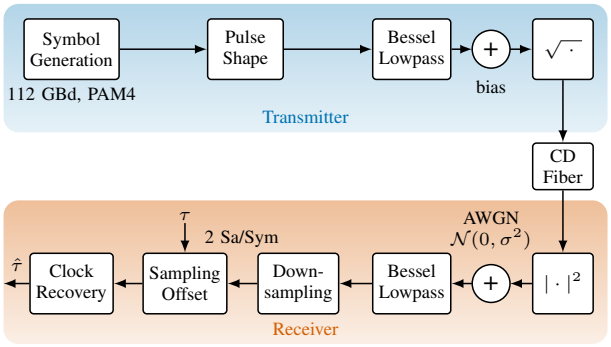

Fig. 1. Simulated IM/DD link for clock recovery algorithm evaluation.

of 78 GHz (70% of symbol rate) to account for bandwidth limitations. For practical transceivers, the component bandwidth can be significantly lower which is examined in section IV. CD is added to the waveform and the signal detection is modeled as square-law operation. Fig. 2(a) shows the transmit and receive spectrum for 2^{12} symbols, 24 ps/nm accumulated CD, and averaged over 100 signal realizations to obtain a smooth averaged spectrum. For a small ER of the modulator of 0.1 dB, the power fading effect is dominating and the spectral dips in accordance to eq. (13) are clearly visible. For a more realistic ER of 10 dB, the nonlinear interference is emphasized in the power fading dips and tones at multiples of the symbol rate are generated. Fig. 2(b) shows the amplitude ratio function in decibel between the transmitted $\underline{\tilde{s}}(f)$ and received $\underline{\tilde{x}}(f)$ signal, where $\underline{\tilde{s}}(f)$ and $\underline{\tilde{x}}(f)$ are the transmit signal after the Bessel lowpass and receive signal after photodetection, respectively, and the power fading function. The amplitude ratio function agrees well with the power fading function. For higher ER, the nonlinear interference increasingly distorts the signal amplitude. In the spectral dips arising from the NRZ pulse at 112 GHz and 224 GHz, the distortion due to nonlinear interference dominates as less signal power is present. Similarly, Fig. 2(c) compares the phase difference between the transmit and receive signals $\arg\{\underline{\tilde{s}}(f)\underline{\tilde{x}}^*(f)\}$ with the phase of the power fading function. In case of dominating power fading (left plot), the negative sign of the power fading frequency response $\tilde{h}_{\mathrm{PF}}(f)$ causes a π-phase shift. In addition, the nonlinear intersymbol interference (ISI) will cause a phase distortion around the spectral nulls, as indicated for the case with a 10 dB ER (right plot).

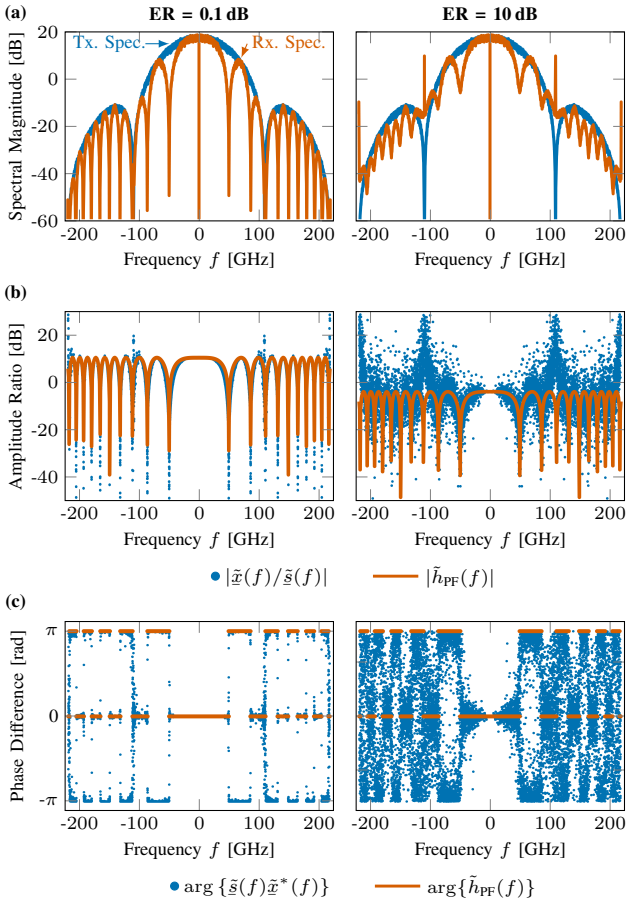

Fig. 2. Simulated effect of CD on the received 112-GBd PAM4 NRZ signal in a direct-detection link with zero sampling offset $\tau = 0$. (a) Transmit (blue) and receive (orange) signal spectrum for an accumulated dispersion of $LD_{\mathrm{CD}} = 24$ ps/nm. Amplitude ratio function in decibel (b) and phase difference (c) between the transmit and receive signal (blue) and power fading function (red).

B. Effect of Chromatic Dispersion on Clock Recovery

To investigate the effect of CD on the clock recovery analytically, we consider a mathematical channel model in FD, where $n = \{1, \ldots, N\}$ is the frequency bin index of an N-point FFT. The zero-frequency is therefore assigned to $n = 1$. For convenience, we consider a clock recovery, which is implemented at twofold oversampling. A twofold upsampled symbol sequence $\tilde{a}_n$ (sample rate upconversion by inserting zeros in between the symbols) is interpolated by a multiplication of the frequency response $\tilde{p}_n$ of a pulse shaping filter. The optical channel is modeled as dispersive medium with a linear response $\tilde{h}_{\mathrm{PF},n}$ and nonlinear interference term $\underline{\tilde{\varepsilon}}_n$ (see eq. (9)). Finally, the channel adds additive white Gaussian noise (AWGN) $\underline{\tilde{n}}_n$. At the receiver side, another lowpass filter $\tilde{g}_n$ is emulating bandwidth limitations and hence suppresses out-of-band noise. The lowpass filters $\tilde{p}_n, \tilde{g}_n \geq 0$ are assumed to be real-valued and non-negative in FD since any nonlinear spectral phase can be pre-compensated in the DSP. Lastly, we consider a sampling offset τ between the transmit and receive clocks which is normalized to the symbol period and can be assumed to be constant within a block of N samples. The n-th element $\underline{\tilde{x}}_n$ of the N-point FFT of the received signal can then be expressed as

$$\underline{\tilde{x}}_n = \left(\underline{\tilde{a}}_n \tilde{p}_n \tilde{h}_{\mathrm{PF},n} + \underline{\tilde{\varepsilon}}_n + \underline{\tilde{n}}_n\right) \tilde{g}_n \, \mathrm{e}^{\mathrm{j}\, 4\pi \frac{n}{N} \tau} \,. \tag{14}$$

Most common non-data-aided clock recoveries compute the spectral autocorrelation to generate a so-called clock tone at symbol rate, whose phase is proportional to the sampling offset [17], e.g., the algorithms proposed in [18], [19], [20], [21], [22], [23], [24]. The spectral autocorrelation at symbol rate

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

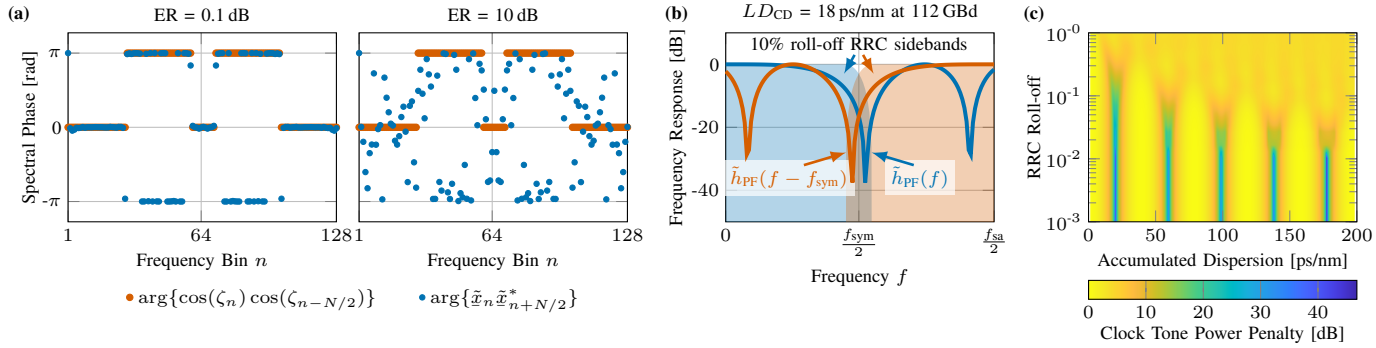

Fig. 3. (a) Phase of cosine product $\arg\{\cos(\zeta_n)\cos(\zeta_{n-N/2})\}$ (red) and signal phase $\arg\{\underline{\tilde{x}}_n\underline{\tilde{x}}^*_{n+N/2}\}$ relevant for the timing estimation (blue) over the frequency bins n for an ER of 0.1 and 10 dB. The received signal is simulated without noise and is distorted by 24 ps/nm accumulated CD and a 256-point FFT is computed. (b) Left and right sideband of a 10% RRC-shaped spectrum indicated by the blue- and red-shaded areas and power fading function in decibel. (c) Clock tone power penalty over accumulated dispersion and RRC roll-off factors.

$R_{\tilde{x}\tilde{x}} = 2/N \sum_{n=1}^{N/2} \underline{\tilde{x}}_n \underline{\tilde{x}}^*_{n+\frac{N}{2}}$ can be expressed as

$$\begin{aligned} R_{\tilde{x}\tilde{x}} = \frac{2}{N}\sum_{n=1}^{N/2} &\left(\underline{\tilde{a}}_n \tilde{p}_n \tilde{h}_{\mathrm{PF},n} + \underline{\tilde{\varepsilon}}_n + \underline{\tilde{n}}_n\right) \\ &\times \left(\underline{\tilde{a}}^*_{n+\frac{N}{2}} \tilde{p}_{n+\frac{N}{2}} \tilde{h}_{\mathrm{PF},n+\frac{N}{2}} + \underline{\tilde{\varepsilon}}^*_{n+\frac{N}{2}} + \underline{\tilde{n}}^*_{n+\frac{N}{2}}\right) \\ &\times \tilde{g}_n \tilde{g}_{n+\frac{N}{2}}\, \mathrm{e}^{\mathrm{j}\,2\pi\tau}\ . \end{aligned} \tag{15}$$

Note that the linear spectral phase of $\tilde{x}_{n+N/2}$, which is caused by the sampling offset, is $\exp(\mathrm{j}\,4\pi(n - N/2)/N\tau)$, since the frequency range from $N/2+1$ to N represents the left sideband at negative baseband frequencies. For convenience, we substitute the pulse shape product and receiver-side low-pass product as a real-valued variable $\tilde{\rho}_n = \tilde{p}_n\tilde{p}_{n+N/2}$ and $\tilde{\gamma}_n = \tilde{g}_n\tilde{g}_{n+N/2}$, respectively. Furthermore, we summarize all noise and nonlinear terms including the nonlinear interference $\underline{\tilde{\varepsilon}}_n$ by $\underline{\tilde{\chi}}_n$. The spectral correlation can then be written as

$$R_{\tilde{x}\tilde{x}} = \frac{2}{N}\,\mathrm{e}^{\mathrm{j}\,2\pi\tau} \sum_{n=1}^{N/2} |\underline{\tilde{a}}_n|^2 \tilde{\gamma}_n \tilde{\rho}_n\, \tilde{h}_{\mathrm{PF},n}\tilde{h}_{\mathrm{PF},n+\frac{N}{2}} + \underline{\tilde{\chi}}_n\,. \tag{16}$$

Due to the cyclostationarity of the transmit symbol sequence, the Fourier transform of the random sequence $\underline{\tilde{a}}_n$ is periodic with symbol rate $N/2$, and thus the product $\underline{\tilde{a}}_n\underline{\tilde{a}}^*_{n+N/2} = \underline{\tilde{a}}_n\underline{\tilde{a}}^*_n = |\underline{\tilde{a}}_n|^2$ results in a real value. The phase of the clock tone then gives an estimation of the sampling offset $\hat{\tau}$ as

$$\begin{aligned} \hat{\tau} &= \frac{1}{2\pi}\arg\{R_{\tilde{x}\tilde{x}}\} \\ &= \frac{1}{2\pi}\arg\left\{\frac{2}{N}\,\mathrm{e}^{\mathrm{j}\,2\pi\tau}\right. \\ &\quad \left.\times \sum_{n=1}^{N/2} |\underline{\tilde{a}}_n|^2 \tilde{\gamma}_n\tilde{\rho}_n \cos(\zeta_n)\cos\left(\zeta_{n-\frac{N}{2}}\right) + \underline{\tilde{\chi}}_n\right\}, \end{aligned} \tag{17}$$

where $\zeta_n = \zeta(f = nf_{\mathrm{sa}}/N)$ and f_{sa} is the sampling rate. For the case of no CD, i.e., $\tilde{h}_{\mathrm{PF},n} = 1$ and $\underline{\tilde{\chi}}_n = 0$, the phase of the clock tone is only proportional to the sampling offset τ. However, in the presence of CD, the timing estimation is impaired by the power fading and the nonlinear interference contained in $\underline{\tilde{\chi}}_n$.

To investigate the effect of CD in simulation, we simulate a 112-GBd IM/DD system, which is impaired by CD and AWGN, as displayed in Fig. 1. At the transmitter, a sequence of PAM4 symbols is generated. Afterwards, each symbol is repeated by eight to model NRZ pulses with a rise time of zero or the symbol sequence is upsampled by inserting zeros between the symbols and interpolated using an RRC pulse-shaping filter. Bandwidth limitations of a digital-to-analog converter (DAC) and other analog components at the transmitter are taken into account by applying a 5th-order Bessel lowpass filter with a 3-dB bandwidth of 78 GHz. To model a non-negative optical power, a bias is added and the optical field amplitude is obtained by the square-root of the optical power. To simulate the fiber-optic channel of length L, the resulting signal is impaired by CD for a certain CD coefficient D_{CD}. At the receiver side, the signal is detected by a photodiode, which is modeled as square-law detector and AWGN is added to set a certain electrical signal-to-noise ratio (SNR) that is normalized to the symbol rate. The resulting signal is once again filtered by a Bessel lowpass and downsampled to 2 Sa/Sym. Finally, a sampling offset τ is added before the clock recovery.

The influence of CD on the clock recovery performance can be observed through three effects. The first effect is depicted in Fig. 3(a), which shows the phase of the cosine product $\arg\{\cos(\zeta_n)\cos(\zeta_{n-N/2})\}$, which can be either 0 or π, in orange and the phase of $\arg\{\underline{\tilde{x}}_n\underline{\tilde{x}}^*_{n+N/2}\}$ for each frequency bin n in blue. To see the effect of CD, we add no AWGN and neglect the influence of the sampling offset on the FD phase for the moment by setting $\tau = 0$. The FFT size is $N = 256$ (128 symbols) and the accumulated dispersion is 18 ps/nm at 112 GBd. Since $\tau = 0$, the phase difference should be zero for each frequency bin. However, there is a phase shift of π for the frequency bins where the product of $\cos(\zeta_n)$ and $\cos(\zeta_{n+N/2})$ have opposite sign. Furthermore, the nonlinear interference in $\underline{\tilde{\chi}}_n$ distorts the phase which becomes especially relevant in the dips of the power fading function $\tilde{h}_{\mathrm{PF}}(f)$. The weighted sum over the cosine product in eq. (17) then has either positive or negative sign. A negative sign leads to a π-phase shift, which results in a false estimation of the sampling offset by half a

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

symbol period and therefore the closed eye is interpreted as the ideal sampling point. However, this constant sampling offset can be compensated manually or by an adaptive equalizer.

The second effect occurs for certain dispersions, where the weighted sum of the cosine product approaches zero. In this case, the left term of the sum in eq. (17) vanishes and the nonlinear interference $\tilde{\chi}_n$ dominates the phase estimation and leads to a inconsistent clock phase estimation. This effect can be seen for the modified version of the algorithm by Barton and Al-Jalili [23] (see eq. (21)) at 11 ps/nm in Fig. 4(b).

The third effect occurs for narrow-band lowpass responses $\tilde{p}$ and $\tilde{g}$, i.e., for strong bandwidth limitations (FTN signals) or low RRC roll-off signals, and is visualized in Fig. 3(b) by showing the overlapping left and right sidebands of a 10% roll-off RRC frequency response (blue and red area) and the power fading function (blue and red curve) in decibel at 112 GBd and 18 ps/nm accumulated dispersion. The term proportional to the cosine product in eq. (17) disappears if the nulls of the cosines fall within the narrow region of the overlapping sidebands. In Fig. 3(c), we compute the clock tone power penalty (CTPP) in decibel by comparing the clock tone resulting from the RRC frequency response with the one impaired by CD as

$$\text{CTPP} = 10\log_{10}\left\{\frac{\sum_{n=1}^{N/2}\tilde{\rho}_n\cos\left(\zeta_n\right)\cos\left(\zeta_{n-\frac{N}{2}}\right)}{\sum_{n=1}^{N/2}\tilde{\rho}_n}\right\}. \quad (18)$$

For a roll-off of 1, the clock tone power penalty is about 3.8 dB for an accumulated dispersion larger than 5 ps/nm, while for low roll-offs the periodic extinction of the clock tone can be observed for dispersion values, where the power fading dips fall around half the symbol rate, i.e., where $|f_m| = f_{\text{sym}}/2$ is valid for eq. (13). Hence, eq. (13) can be solved for $f_{\text{sym}}/2$ and the affected accumulated dispersion values can be found as

$$LD_{\text{CD}} = \frac{2(1+2m)c}{f_{\text{sym}}^2\lambda^2}. \quad (19)$$

The three effects repeat periodically as the accumulated CD increases, since the function $\tilde{h}_{\text{PF}}(f)$ is periodic as well. It shall be noted that the effects occur for timing error detectors used in feedback (FB) control loops and timing estimators used in feedforward clock recovery architectures. Furthermore, this effect can be observed regardless of whether the timing is estimated in time or frequency domain.

III. Chromatic Dispersion Tolerant Clock Recovery for Non-Return-to-Zero Signals

As a consequence of CD on the clock recovery for data communication, the sample period offset at twofold oversampling now results in sampling at the signal transitions, leading to strong additional ISI. This effect can either be compensated by a constant half-symbol-period sampling offset using an adaptive equalizer after the clock recovery or by manual addition of a π-phase shift for the affected fiber lengths. However, both approaches will still suffer by the clock recovery failure due to a vanishing clock tone if the summed cosine product is close to zero (second effect). For this reason, we propose a modification that makes the FD clock recovery algorithm robust to CD in direct-detection systems.

The starting point of our consideration is the algorithm of Barton and Al-Jalili (BAJ) [23], where a timing estimate is computed from two frequency components at bin n and $n + N/2$ as

$$\hat{\tau}_{\text{BAJ}} = \frac{1}{2\pi}\arg\{\underline{\tilde{x}}_n\underline{\tilde{x}}^*_{n+N/2}\}. \quad (20)$$

In [24], [25], [26], the BAJ algorithm is modified by averaging the product $\underline{\tilde{x}}_n\underline{\tilde{x}}^*_{n+N/2}$ over the frequency axis to reduce the phase estimation inaccuracy caused by random noise, resulting in the autocorrelation (see eq. (17))

$$\hat{\tau}_{\text{mod-BAJ}} = \frac{1}{2\pi}\arg\left\{\sum_{n=1}^{N/2}\underline{\tilde{x}}_n\underline{\tilde{x}}^*_{n+\frac{N}{2}}\right\}. \quad (21)$$

At this point we would like to point out that for power-efficient hardware implementation of these algorithms for low-power IM/DD systems, the positions of the overlapping areas of the sidebands can be taken into account in order to implement the algorithms for fractional oversampling [27].

Clock recovery in FD offers the major advantage that we can easily compensate for the cosine terms. In order to mitigate the effect of CD, we first correct the Fourier transform $\underline{\tilde{x}}$ for the phase rotation of the cosine product

$$\underline{\tilde{y}}_n = \underline{\tilde{x}}_n\,\text{sgn}\left\{\cos\left(\zeta_n\right)\right\} \quad (22)$$

and then apply the timing estimator as

$$\hat{\tau}_{\text{CD-BAJ}} = \frac{1}{2\pi}\arg\left\{\sum_{n=1}^{N/2}\underline{\tilde{y}}_n\underline{\tilde{y}}^*_{n+\frac{N}{2}}\right\}. \quad (23)$$

Doing so, the power fading function will always be positive and the spectral correlation will add up constructively, hence, it avoids the π-phase shift (first effect) and the weighted cosine to approach zero (second effect). This modification requires the information of the total accumulated dispersion of the fiber link, for example from the CD estimation performed in the link configuration. From this, $\tilde{h}_{\text{PF},n}$ can be pre-computed as a reference model. Note that small CD measurement inaccuracies may lead to frequency components in $\underline{\tilde{y}}_n$ which have an incorrect sign-correction. However, this effect is mitigated due to the correlation averaging over $N/2$ frequency bins. Furthermore, it should be emphasized that this modification only marginally increases the computational effort compared to the mod-BAJ, as only the sign of the Fourier transform has to be adjusted. Fig. 4 shows the simulation results for a 112-GBd PAM4 signal with NRZ pulse shape and 78 GHz 6-dB bandwidth (one Bessel filter on, both, the Tx and Rx side) for various accumulated CD values. A common performance metric to evaluate clock recovery algorithms is the jitter, defined as the variance of the deviation of the estimated clock phase from the actual clock phase in decibels as $20\log_{10}(\text{std}(\hat{\tau}-\tau))$. However, if the accumulated CD leads to a π-phase shift, it is possible that a timing estimate with low uncertainty is obtained, which however leads to a high jitter since $\hat{\tau}-\tau$ is large. For this reason, we visualize the performance in Fig. 4(a) by plotting the timing estimate error

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

$\hat{\tau} - \tau$ over the accumulated CD and keep $\tau = 0$ in simulation. Since the π-phase shift leads to phase jumps due to phase wrapping, we plot the absolute value of the estimation error for a simpler overview. To recognize the effect of the CD, we add no AWGN to the signal for now. For each CD value, we simulate 100 waveforms and calculate the timing estimate for a block length of $N = 256$ samples. We plot the mean estimation error (dark blue curve) and the $\pm\sigma$ confidence interval of the estimation error as the light blue area around it. Furthermore, we indicate the areas in which the cosine product including the pulse shape leads to a change of the sign in gray. For the mod-BAJ algorithm with 0.1 dB ER, the phase jumps can be clearly noticed. Around those phase jumps, the cosine product is close to zero and the timing estimate is distorted by the nonlinear interference, which leads to an increase of the standard deviation. This effect is emphasized for an ER of 10 dB. With our proposed CD-BAJ algorithm, this can be counteracted so that no phase jumps occur and the standard deviation is low. Only a constant estimation error linearly depending on the accumulated CD cannot be corrected, but can easily be compensated for by an consecutive adaptive equalizer. Fig. 4(b) shows the timing estimate over 50,000 simulated sampling offsets for different CD values and an SNR of 20 dB for the mod-BAJ and the CD-tolerant algorithm. The colored symbols mark the position in Fig. 4(a).

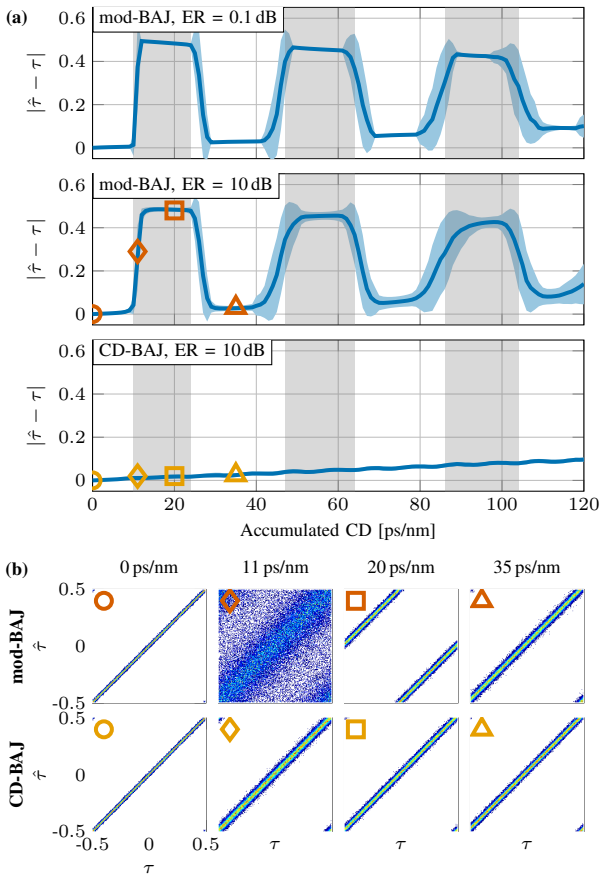

Fig. 4. Clock recovery performance for a 112-GBd PAM4 signal with NRZ pulse shape and 78-GHz 6-dB bandwidth limitation. (a) Absolute mean error of the timing estimate (dark blue line) with its $\pm\sigma$ confidence interval indicated by the light blue area around it computed from 100 waveforms over accumulated CD and no AWGN. (b) Exemplary linear curves of the timing estimate over 50,000 simulated sampling offsets using the mod-BAJ and the CD-tolerant BAJ algorithm for 20 dB SNR. The inset symbols mark the position in subfigure (a).

IV. Chromatic Dispersion Tolerant Clock Recovery for Low-Roll-Off and Faster-Than-Nyquist Signals

The presented CD-tolerant BAJ algorithm only compensates for the phase rotation caused by the cosine and the clock tone extinction caused by counter-phase cosine products. However, in case of low-roll-off signals or extreme bandwidth limitations, this modification does not compensate for the power fading of the overlapping areas of the signal sidebands. In such a case, clock recovery initially designed for FTN signals, i.e., for signals whose 3-dB bandwidth are significantly lower than half the symbol rate, may be beneficial. In order to broaden the spectrum of such signals, thus allowing to recover a clock tone, a nonlinear operation is commonly used, for example the magnitude squared in time domain, which is also referred to as 4-th power method [28], [29], [30]. Similar to the algorithm in [29], the squaring in time domain $z_k = |x_k|^2 = x_k x_k^*$ can also be described as a convolution in FD as

$$\tilde{\underline{z}}_n = \sum_{m=1}^{N} \tilde{\underline{x}}_m \tilde{\underline{x}}^*_{m-n+1} \,. \tag{24}$$

Note that the complex conjugate time domain signal is considered and thus the algorithm can also be applied to complex-valued modulation formats. Although PAM signals are real-valued, the complex conjugation in the FD must be considered in order to obtain the power correlation. At this point, we would like to highlight that under severe bandwidth limitations (e.g., FTN signaling), the analog-to-digital converter (ADC) is capable of capturing the full signal spectrum even with baud-rate sampling, hence, enabling significant power savings. Following the initial sampling, digital resampling can be employed to facilitate oversampled digital clock recovery. Accordingly, the 4-th power version of the mod-BAJ algorithm results in

$$\begin{aligned}\hat{\tau}_{\text{mod-4P-BAJ}} &= \frac{1}{2\pi} \arg\left\{ -\sum_{n=1}^{N/2} \tilde{\underline{z}}_n \tilde{\underline{z}}^*_{n+\frac{N}{2}} \right\} \\ &= \frac{1}{2\pi} \arg\left\{ -\sum_{n=1}^{N/2} \left[\sum_{m=1}^{N} \tilde{\underline{x}}_m \tilde{\underline{x}}^*_{m-n+1} \right] \right. \\ &\quad \left. \times \left[\sum_{m=1}^{N} \tilde{\underline{x}}_m \tilde{\underline{x}}^*_{m+\frac{N}{2}-n+1} \right]^* \right\}.\end{aligned} \tag{25}$$

Finally, we include the power fading correction from eq. (22) to obtain the CD-tolerant version of the 4th-power BAJ timing estimator as

$$\begin{aligned}\hat{\tau}_{\text{CD-4P-BAJ}} &= \frac{1}{2\pi} \arg\left\{ -\sum_{n=1}^{N/2} \left[\sum_{m=1}^{N} \tilde{\underline{y}}_m \tilde{\underline{y}}^*_{m-n+1} \right] \right. \\ &\quad \left. \times \left[\sum_{m=1}^{N} \tilde{\underline{y}}_m \tilde{\underline{y}}^*_{m+\frac{N}{2}-n+1} \right]^* \right\}.\end{aligned} \tag{26}$$

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

In analogy to the CD-BAJ in Fig. 4(a), Fig. 5(a) shows the absolute mean estimation error including the error's $\pm\sigma$ confidence interval over the accumulated CD for the CD-BAJ, mod-4P-BAJ, and the CD-tolerant version CD-4P-BAJ for a 112-GBd PAM4 signal with a 2% RRC roll-off. For the CD-BAJ, we calculate the spectral correlation only for the non-zero frequency components, which is given by the RRC roll-off. For a roll-off of 0.02, the spectral correlation is calculated for only 3 frequency components [24]. Due to the more narrow signal bandwidth, the sign changes of the cosine product are correspondingly less frequent and are again indicated by the gray area. Around these positions, the narrow overlap of the signal spectrum is canceled out by the power fading, which is why the CD-BAJ produces incorrect timing estimates with high standard deviation. While the mod-4P-BAJ works for extremely bandwidth-limited signals, the power fading in particular leads to an incorrect and highly uncertain timing estimate over the entire range of all accumulated CD values. By first compensating the sign of the power fading again, the CD-4P-BAJ algorithm becomes much more robust against CD. The improved robustness of the CD-4P-BAJ comes at the price of increased computational effort compared to the CD-BAJ. The question therefore arises when it performs better (except for the case when clock tone extinction occurs at around 20, 60, and 100 ps/nm). For this purpose, we simulated the jitter and swept the RRC roll-off and the Bessel filter bandwidth for the CD-BAJ and CD-4P-BAJ in Fig. 5(b) for a constant accumulated CD of 0 ps/nm and 30 ps/nm and an SNR of 20 dB. When comparing the jitter between 0 and 30 ps/nm, a penalty due to the nonlinear interference terms can be observed. The CD-4P-BAJ outperforms the CD-BAJ for roll-off less than 0.06 while the CD-4P-BAJ performance rapidly degrades for roll-offs larger than about 0.08, due to aliasing caused by the squaring in time domain. Sweeping the 6-dB bandwidth at a 0.02 roll-off reveals the superior performance of the 4-th power method compared to the mod-BAJ algorithm. The penalty caused by nonlinear interferences can be reduced by additional averaging over consecutive timing estimate realizations by buffering and averaging the complex values before calculating the angle [31]. With a moving average over 8 clock tones, the jitter can be significantly reduced, here shown by the dashed lines. The CD-4P-BAJ then performs better for roll-offs smaller than 0.05.

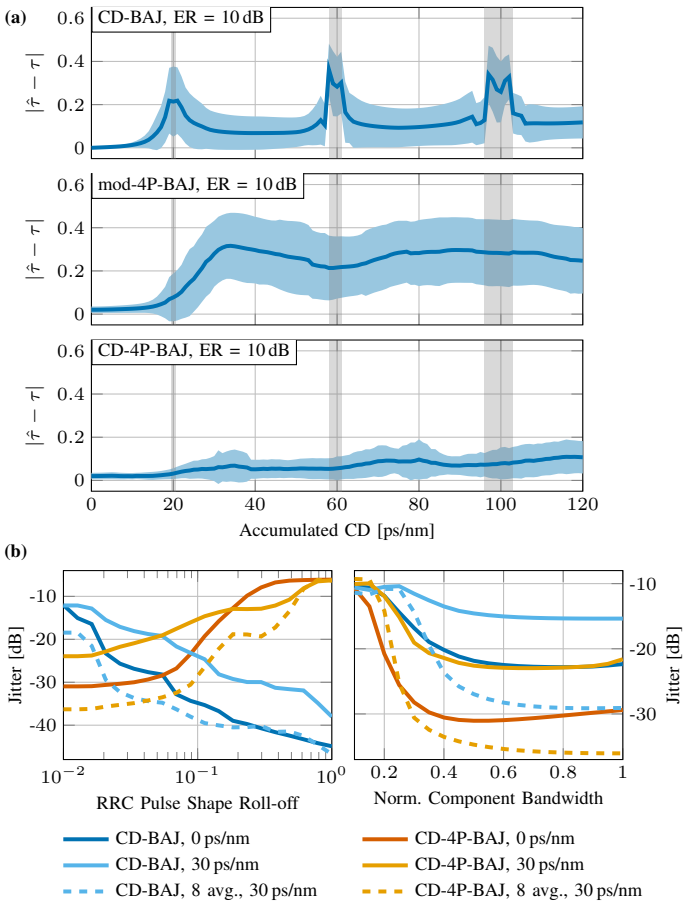

Fig. 5. Clock recovery performance for a 112-GBd PAM4 signal with 2% roll-off RRC pulse shape and 78-GHz 6-dB bandwidth limitation. (a) Absolute mean timing estimate error (dark blue line) with its $\pm\sigma$ confidence interval indicated by the light blue area around it computed from 100 waveforms over accumulated CD and without AWGN. (b) Jitter over RRC roll-off and component bandwidth normalized to the symbol rate computed from 50,000 waveforms with 20 dB SNR as well as 0 ps/nm and 30 ps/nm accumulated CD. The dashed lines show the jitter when an 8-tap moving average filter is applied to the clock tone.

V. Experimental Validation

In this section, we experimentally validate our model for the effect of CD on clock recovery and our proposed algorithms. For this purpose, we set up an IM/DD system as depicted in Fig. 6. At the transmitter and receiver sides, we use the Keysight USPA real-time field-programmable gate array (FPGA) platform, equipped with a DAC and an ADC (formerly Micram DAC3 and ADC3), respectively. We generate a 34-GBd PAM4 signal (limited by available hardware) using the FPGA with either an NRZ or an RRC pulse shape with 0.1 roll-off. The DAC has a nominal 6-dB bandwidth of 28 GHz and is sampling at 68 GSa/s, thus, NRZ pulses are generated by simply repeating each symbol twice. The amplified analog signal is then used to modulate the light emitted by an external-cavity laser (ECL) at 1550 nm utilizing a Mach-Zehnder modulator (MZM) with a 3-dB bandwidth of 25 GHz and an ER of 22 dB according to the datasheet. To realize various values of accumulated dispersion, we use a combination of 10.56 km-long fiber spools combined with a wavelength-selective switch (WSS), which allows fine-tuning within $\pm$100 ps/nm. The fiber spools have an estimated CD coefficient of 16.3 ps/nm/km, i.e., a single spool, a concatenation of two spools, and a concatenation of three spools features a total accumulated dispersion of 172 ps/nm, 344 ps/nm, and 516 ps/nm, respectively. Therefore, the dispersion can be set in a range from -100 to 616 ps/nm. The optical power at the output of the WSS is 3.6 dBm before the fiber spools. Finally, a variable optical attenuator (VOA) allows to set a constant received optical power of -2 dBm for 10.56 km and 21.12 km fiber distance and -4.5 dBm for 31.68 km. At the receiver side, a 27-GHz 3-dB bandwidth optical detector (Optilab PR-40G-M) consisting of a photodiode and a transimpedance amplifier (TIA) is used to convert the optical signal into a voltage.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

Finally, the ADC with nominal 3-dB bandwidth of about 37 GHz is sampling the received signal at 68 GSa/s, which is is stored in the receiver FPGA for offline processing.

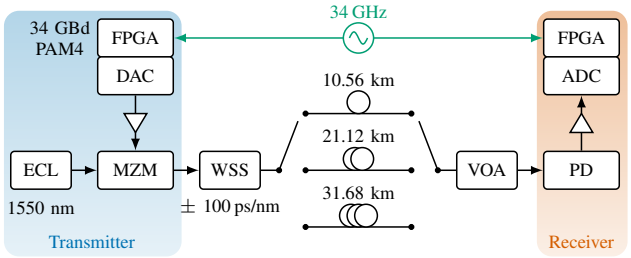

Fig. 6. Experimental IM/DD setup to investigate the impact of accumulated dispersion on the clock recovery. The amount of accumulated dispersion is set by using different fiber spools together with a WSS.

Since the group delay of the channel varies with different dispersion settings applied at the WSS during the experiment, the corresponding constant sampling phase estimation also differs for each dispersion value. Therefore, we evaluate the timing estimates separately for selected characteristic dispersion levels. To this end, the simulated timing estimates as a function of accumulated CD for the mod-BAJ and the CD-BAJ for 34 GBd NRZ signals considering a 25-GHz 3-dB bandwidth Bessel lowpass filter are shown in Fig. 7(a). Similarly, we compare the CD-BAJ and CD-4P-BAJ for RRC-shaped signals with 10% roll-off in Fig. 8(a). We then select certain characteristic accumulated dispersion values (indicated by the orange symbols in Fig. 7(a) and Fig. 8(a)) and show the timing estimates over time obtained from the experimental waveforms in Fig. 7(b) and Fig. 8(b). In offline processing, we split the received waveforms from the experiment into blocks of $N = 256$ samples. Afterwards, the timing estimate is computed for each of these blocks using the clock recovery algorithms described above. Due to the strong bandwidth limitation, it is sufficient to consider only a part of the frequency components in the summation of the correlation (eq. (21) and eq. (23)) [27] for the NRZ signal. For this reason, we use only the center $N/4$ frequency components in the summation over n in order to reduce the number of frequency components outside the signal spectrum, which primarily contain noise and nonlinear interference terms, and to reduce the computational complexity. We compare the timing estimation for four dispersion values by means of conventional clock recovery using the mod-BAJ and using the proposed CD-tolerant BAJ algorithm in Fig. 7(b) for the NRZ signal. The experimental results are in very good agreement with our theoretical considerations and the simulation in Fig. 7(a). For 195 ps/nm, indicated by a circle, the π-phase shift can be recognized at low estimation deviation. For about 308 ps/nm, indicated by a diamond, the sum over the cosine products is approaching zero, leading to a cancellation of the clock tone and, therefore, nonlinear interference terms are more pronounced. The conventional clock recovery fails, while the proposed CD-tolerant algorithm still provides a timing estimate with low deviation. At 403 ps/nm is a region with no π-phase shift, where both algorithms provide identical estimations. Finally, at 500 ps/nm the clock tone extinction starts again, but the CD-tolerant BAJ continues to perform well.

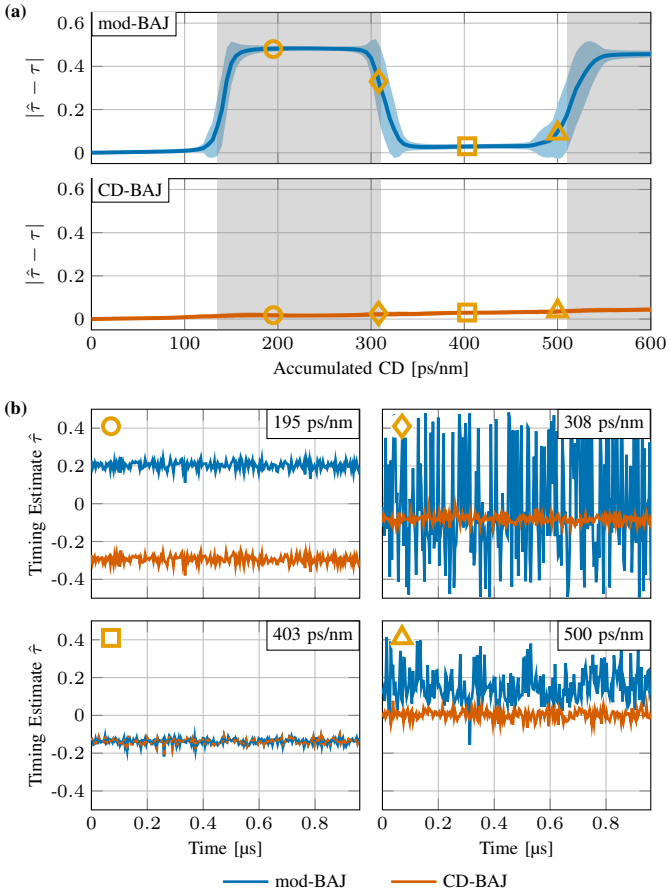

Fig. 7. Timing estimation results for 34-GBd PAM4 NRZ signals. (a) Simulated timing estimate as a function of accumulated CD obtained using the mod-BAJ (blue) and CD-BAJ (red) algorithms. Characteristic dispersion values are indicated by yellow markers. (b) Timing estimates obtained by offline processing of recorded waveforms from the experiment for the four characteristic dispersion values indicated in (a). For a block size of $N = 256$ samples, $N/4$ frequency bins per sideband are evaluated for the mod-BAJ (red line) and CD-BAJ algorithm (blue line), respectively.

Next, we consider the experimental clock recovery performance for an RRC-shaped signal with 0.1 roll-off to observe the third effect mentioned in section II-B. Since the signal spectrum is now more narrow, it is sufficient to sum only over the non-zero frequency components, i.e., where $\tilde{\gamma}_n \tilde{\rho}_n > 0$, for the CD-BAJ algorithm. In this case, we sum over 12 out of 128 frequency components per sideband. As aliasing starts to decrease the performance of the 4-th power method, only the middle $N/4$ bins in the outer sum over index n in eq. (26) are evaluated. Since fewer bins are used, the influence of noise on the estimate is more significant. For this reason, an additional 8-tap moving average of the spectral correlation is used. As the frequency range of the overlapping RRC-shaped sidebands is small for such a signal, a π-phase shift is less frequent, but a clock tone extinction due to the spectral nulls caused by power fading is possible. For a 34-GBd signal, this can be observed in simulation in Fig. 8(a) at around 200 ps/nm by an increased uncertainty of the timing estimate. This observation is also confirmed in the experiment

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

when considering the timing estimate for 195 ps/nm using the CD-BAJ and the CD-4P-BAJ. The proposed CD-BAJ shows a high standard deviation due to the few/no frequency bins over which it can average, while the CD-4P-BAJ exhibits a significantly lower standard deviation and is therefore tolerant to, both, the π-phase shift and the clock tone extinction due to spectral dips falling in the overlap region. At 322 ps/nm there are enough frequency bins available for the CD-BAJ to estimate the timing with low uncertainty. However, the CD-4P-BAJ shows a higher uncertainty caused by aliasing, showing the price to pay when using this algorithm. This effect is also reflected in Fig. 8(b), where the CD-4P-BAJ shows more than 10 dB worse jitter compared to the CD-BAJ at a roll-off of 0.1. Accordingly, the CD-4P-BAJ is only advantageous in the case of strong bandwidth limitation, either caused by a narrowing of the signal spectrum when the spectral dips fall directly on the edges of the signal spectrum, or in the case of extremely low roll-off factors or FTN signals.

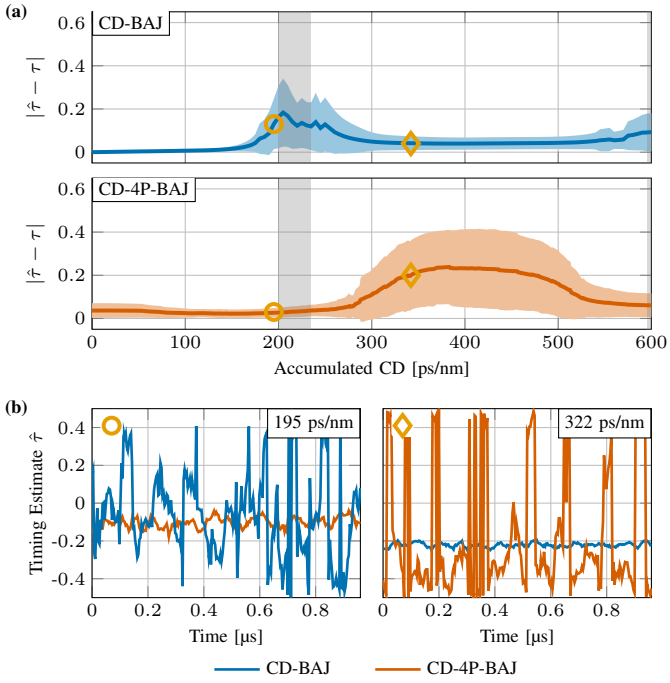

Fig. 8. Timing estimation results for 34-GBd PAM4 signals exhibiting a RRC pulse shape with 0.1 roll-off. (a) Simulated timing estimate as a function of accumulated CD obtained using the CD-BAJ (blue) and CD-4P-BAJ (red) algorithms. Characteristic dispersion values are indicated by yellow markers. (b) Timing estimates obtained by offline processing of recorded waveforms from the experiment for the two characteristic dispersion values indicated in (a). For a block size of $N = 256$ samples and 8-tap MA of the spectral correlation, 12 frequency bins per sideband are evaluated for the CD-BAJ (red line) and $N/4$ frequency bins per sideband for the CD-4P-BAJ algorithm (blue line), respectively.

VI. Conclusion

This work analytically examines the influence of CD on clock recovery in direct-detection systems. Power fading is identified as the dominant impairment, causing π-phase errors at specific dispersion values. Additionally, the clock tone may cancel out when the weighted cosine product approaches zero or when spectral narrowing due to power fading occurs. These effects are consistently observed in both simulation and experiment. To mitigate them, CD-tolerant clock recovery algorithms for NRZ and RRC/FTN signals are proposed, which compensate phase distortions in the frequency domain. Reliable clock recovery under severe CD is critical for the convergence of adaptive equalizers. This work therefore makes an important contribution to improving the robustness of future high-speed direct-detection systems equipped with DSP. Future research may consider low-power baud-rate clock recovery and the investigation of the additional influence of modulator chirp.

References

[1] M. M. H. Adib, P. Matalla, C. Füllner, S. Li, E. Giacoumidis, C. Raack, U. Menne, M. Straub, M. Saier, C. Schweickert, S. Orf, M. Gontscharow, T. Käfer, M. Färber, A. Richter, R. Bonk, and S. Randel, "Optical access networks for smart sustainable cities: From network architecture to fiber deployment," *J. Opt. Commun. Netw.*, 2025.

[2] P. Maniotis and D. M. Kuchta, "Exploring the benefits of using co-packaged optics in data center and AI supercomputer networks: a simulation-based analysis [invited]," *J. Opt. Commun. Netw.*, vol. 16, no. 2, pp. A143–A156, Feb. 2024.

[3] S. J. B. Yoo, "New trends in photonic switching and optical networking architectures for data centers and computing systems [invited]," *J. Opt. Commun. Netw.*, vol. 15, no. 8, pp. C288–C298, Aug. 2023.

[4] S. T. Le, G. De Valicourt, P. Pupalaikis, R. Giles, M. Lamponi, L. Elsinger, S. Liu, B. Sawyer, J. Proesel, E. Ho, K. Liu, G. Homsey, J. Lopez, Z. Zhu, S. Corteselli, L. Alloin, C. Daunt, M. Ferriss, B. Rahmani, F. Warning, A. Bruno, S. Abbaslou, M. Zaman, Z. Pan, G. Fischer, P. Haigh, G. Reichert, A. Gazman, F. Fesharaki, and P. Winzer, "1.6-Tbps low-power linear-drive high-density optical interface (HDI/O) for ML/AI," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2024, pp. 1–3.

[5] M. Jacques, Z. Xing, A. Samani, X. Li, E. El-Fiky, S. Alam, O. Carpentier, P.-C. Koh, and D. V. Plant, "Net 212.5 Gbit/s transmission in o-band with a SiP MZM, one driver and linear equalization," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2020, pp. 1–3.

[6] X. Pang, T. Salgals, H. Louchet, D. Che, M. Gruen, Y. Matsui, T. Dippon, R. Schatz, M. Joharifar, B. Krüger, F. Pittala, Y. Fan, A. Udalcovs, L. Zhang, X. Yu, S. Spolitis, V. Bobrovs, S. Popov, and O. Ozolins, "200 Gb/s optical-amplifier-free IM/DD transmissions using a directly modulated o-band DFB+R laser targeting LR applications," *J. Lightw. Technol.*, vol. 41, no. 11, pp. 3635–3641, June 2023.

[7] X. Zhou, C. F. Lam, R. Urata, and H. Liu, "State-of-the-art 800G/1.6T datacom interconnects and outlook for 3.2T," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2023, pp. 1–3.

[8] P. Zhu, Y. Yoshida, K. Akahane, and K. ichi Kitayama, "High-speed reach-extended IM-DD system with low-complexity DSP for 6G fronthaul," *J. Opt. Commun. Netw.*, vol. 16, no. 1, pp. A24–A32, Jan. 2024.

[9] *50-Gigabit-capable passive optical networks (50G-PON): Physical media dependent (PMD) layer specification*, ITU-T Recommendation G.9804.3 Std., Mar. 2024.

[10] R. Bonk and E. Harstead, "The road towards 100G and 200G-passive optical networks," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 2024, paper W3D.4.

[11] N. Stojanovic, F. Karinou, Z. Qiang, and C. Prodaniuc, "Volterra and Wiener equalizers for short-reach 100G PAM-4 applications," *J. Lightw. Technol.*, vol. 35, no. 21, pp. 4583–4594, Nov. 2017.

[12] L. Schmalen, V. Lauinger, J. Ney, N. Wehn, P. Matalla, S. Randel, A. von Bank, and E.-M. Edelmann, "Recent advances on machine learning-aided DSP for short-reach and long-haul optical communications," 2024. [Online]. Available: https://arxiv.org/abs/2411.10101

[13] F. N. Hauske, N. Stojanovic, C. Xie, and M. Chen, "Impact of optical channel distortions to digital timing recovery in digital coherent transmission systems," in *Int. Conf. on Transparent Opt. Networks*, Aug. 2010, paper We.D1.4, pp. 1–4.

[14] D. Wang, M. Qiao, K. Lian, and Z. Li, "CD and PMD effect on cyclostationarity-based timing recovery for optical coherent receivers," *J. Lightw. Technol.*, vol. 41, no. 8, Apr. 2023.

[15] J. Wang and K. Petermann, "Small signal analysis for dispersive optical fiber communication systems," *J. Lightw. Technol.*, vol. 10, no. 1, pp. 96–100, Jan. 1992.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in Journal of Lightwave Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JLT.2025.3600353

[16] M. Chagnon, "Optical communications for short reach," *J. Lightw. Technol.*, vol. 37, no. 8, pp. 1779–1797, Apr. 2019.
[17] L. Huang, D. Wang, A. P. T. Lau, C. Lu, and S. He, "Performance analysis of blind timing phase estimators for digital coherent receivers," *Opt. Express*, vol. 22, no. 6, pp. 6749–6763, Mar. 2014.
[18] D. Godard, "Passband timing recovery in an all-digital modem receiver," *IEEE Trans. Commun.*, vol. 26, no. 5, pp. 517–523, Jan. 1978.
[19] F. Gardner, "A BPSK/QPSK timing-error detector for sampled receivers," *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 423–429, May 1986.
[20] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*. New York, NY, USA: John Wiley & Sons, 1998.
[21] S. J. Lee, "A new non-data-aided feedforward symbol timing estimator using two samples per symbol," *IEEE Commun. Lett.*, vol. 6, no. 5, pp. 205–207, May 2002.
[22] W.-P. Zhu, Y. Yan, M. Ahmad, and M. Swamy, "Feedforward symbol timing recovery technique using two samples per symbol," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 11, pp. 2490–2500, Nov. 2005.
[23] S. Barton and Y. Al-Jalili, "A symbol timing recovery scheme based on spectral redundancy," in *IEE Colloquium on Advanced Modulation and Coding Techniques for Satellite Communications*, Jan. 1992, pp. 3/1–3/6.
[24] P. Matalla, J. Krimmer, L. Schmitz, D. Fang, C. Koos, and S. Randel, "Joint blind clock recovery for space-division multiplexed optical transmission systems," *J. Lightw. Technol.*, 2025.
[25] N. Kaneda, A. B. Leven, and S. Weisser, "Symbol timing recovery in polarization division multiplexed coherent optical transmission system," U.S. Patent US 8,655,191 B2, Feb., 2014.
[26] P. Matalla, M. S. Mahmud, C. Füllner, C. Koos, W. Freude, and S. Randel, "Hardware comparison of feed-forward clock recovery algorithms for optical communications," in *Proc. Opt. Fiber Commun. Conf.*, June 2021, paper Th1A.10, pp. 1–3.
[27] A. Josten, B. Baeuerle, E. Dornbierer, J. Boesser, D. Hillerkuss, and J. Leuthold, "Modified godard timing recovery for non integer oversampling receivers," *Appl. Sci.*, vol. 7, no. 7, June 2017.
[28] M. Yan, Z. Tao, L. Dou, L. Li, Y. Zhao, T. Hoshida, and J. C. Rasmussen, "Digital clock recovery algorithm for Nyquist signal," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2013, pp. 1–3.
[29] K.-T. Wu and H. Sun, "Frequency-domain clock phase detector for nyquist WDM systems," in *Proc. Opt. Fiber Commun. Conf.*, Mar. 2014, paper Th3E.2, pp. 1–3.
[30] N. Stojanovic, B. Mao, and Y. Zhao, "Digital phase detector for Nyquist and faster than Nyquist systems," *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 511–514, Mar. 2014.
[31] P. Matalla, M. S. Mahmud, C. Koos, and S. Randel, "Pilot-free digital clock synchronization for continuous-variable quantum key distribution," in *Proc. Eur. Conf. Opt. Commun.*, Oct. 2023, pp. 1386–1389.

Patrick Matalla (Student Member, IEEE) received his B.Sc. and M.Sc. degrees in electrical engineering from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2018 and 2020, respectively. He joined the Institute of Photonics and Quantum Electronics (IPQ) at KIT in 2020, where he is currently working towards his Ph.D. His research interests include digital signal processing in optical communication systems and the algorithm hardware implementation on field-programmable gate arrays (FPGAs).

Christian Koos received the Ph.D. (Dr.-Ing.) degree in electrical engineering from the University of Karlsruhe, Karlsruhe, Germany, in 2007. He is currently a full Professor with the Karlsruhe Institute of Technology, Karlsruhe, Germany, where he is heading the Institute of Photonics and Quantum Electronics. He has co-founded several start-up companies, such as Vanguard Photonics GmbH, Vanguard Automation GmbH, SilOriX GmbH, and DeepLight SA. From 2008 to 2010, he was affiliated with the Corporate Research and Technology Department of Carl Zeiss AG in Oberkochen, Germany, where he led the technology forecast in the area of nanotechnology. He is the author of more than 140 journal papers and more than 30 patent families. His research interests include silicon photonics and hybrid integration concepts along with the associated applications in high-speed communications, optical sensing and metrology, and ultra-fast photonic-electronic signal processing. He was the recipient of several research awards and prestigious grants, such as the ERC Starting Grant in 2011 and ERC Consolidator Grant in 2017.

Sebastian Randel (Senior Member, IEEE) received his Dr.-Ing. degree in 2005 from Technische Universität Berlin, Germany, for his work on high-speed optical-time-division-multiplexed transmission systems. Since 2016, he has been a full Professor at the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, where he co-heads the Institute of Photonics and Quantum Electronics. From 2005 to 2010, he was a Research Scientist at Siemens Corporate Technology, Munich, Germany, leading research and standardization activities in polymer-optical-fiber communications, visible light communications, and optical access networks. From 2010 to 2016, he served as a Member of Technical Staff at Bell Laboratories, Holmdel, NJ, USA. His current research focuses on designing and implementing power-efficient transceivers for high-speed optical communications, passive optical networks, and wireless links at optical and sub-THz frequencies.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/