



Copyright Notice

©2012 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This document was downloaded from Chalmers Publication Library (<http://publications.lib.chalmers.se/>), where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec. 8.1.9 (<http://www.ieee.org/documents/opsmanual.pdf>)

(Article begins on next page)

Modified Cramér-Rao Bound for Clock Recovery in the Presence of Self-Phase Modulation

A. Serdar Tan, Henk Wymeersch, *Member, IEEE*, Pontus Johannisson, Erik Agrell, Magnus Karlsson, and Peter A. Andrekson, *Fellow, IEEE*

Abstract—We investigate the effect of nonlinear self-phase modulation (SPM) on clock recovery in long-haul coherent optical communication systems with low to moderate baud rates. We consider dual-polarization, multilevel quadrature amplitude modulation, and evaluate conventional timing estimators as well as a modified Cramér-Rao bound. As the latter is hard to compute in the presence of nonlinear channel impairments, we derive an approximate closed-form expression. Contrary to linear channel impairments, the nonlinear impairments can degrade the performance for higher input powers. In addition, the bounds exhibit oscillatory behavior due to the oscillations in the signal spectrum in the presence of SPM.

Index Terms—Modified Cramér-Rao bound, coherent optical communication, timing offset estimation, clock recovery, self-phase modulation, fiber optics.

I. INTRODUCTION

RECEIVERS used in coherent optical communications must compensate for various linear and nonlinear impairments. As we move towards multilevel quadrature amplitude modulation (M-QAM) formats, which require higher input powers to achieve a fixed target bit error rate, the impact of the nonlinear impairments is performance-limiting in dual-polarization systems [2].

For moderate baud rates and large channel spacing, self-phase modulation (SPM) is the main nonlinear phenomenon in fiber optical communication systems [3] and is caused by interaction of the light and the optical fiber. This results in changes of the fiber refractive index and leads to nonlinear phase shifts proportional to the signal power, with concomitant spectral broadening. The effect of SPM accumulates since each optical amplifier restores the initial power level. In addition, optical amplifiers introduce amplified spontaneous emission (ASE) noise, leading to nonlinear phase noise (NLPN) [4], which is an interaction between the signal and the noise. Thus, for input power levels above a certain threshold, the nonlinear impairment may become the main limitation in long-haul systems with many amplifier stages. This is in stark

Part of this work was presented in [1], at the 2010 European Conference and Exhibition on Optical Communication.

A. Serdar Tan was with the Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology, and is now with Turk Telekom, Turkey. Email: ahmetserdar.tan@turktelekom.com.tr. Henk Wymeersch and Erik Agrell are with the Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology, SE-412 96 Göteborg, Sweden. Pontus Johannisson, Peter A. Andrekson, and Magnus Karlsson are with the Photonics Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE-412 96 Göteborg, Sweden. Email: {henkw, pontus.johannisson, agrell, magnus.karlsson, peter.andrekson}@chalmers.se.

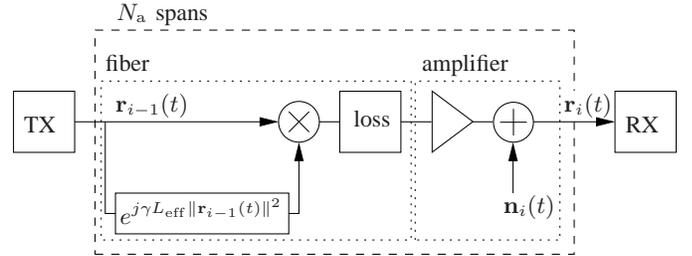


Figure 1. Transmission and receiver model: the signal passes through N_a spans of fiber, the fiber loss is compensated by the amplifier, which also induces ASE noise. In every fiber span, the signal undergoes SPM.

contrast to linear transmission systems, where an increase in input power cannot lead to a decrease in performance.

There exist various studies on dealing with SPM and NLPN, such as [5]–[9], which generally assume perfect clock recovery. As clock recovery typically occurs early in the receiver chain, this assumption may not hold. A useful tool in analyzing the validity of this assumption is through theoretical performance bounds, such as the Cramér-Rao bound (CRB) [10]. The CRB, and variations thereof, provide a fundamental lower bound on the variance of estimators (including clock recovery algorithms).

In this paper, we build on our initial findings from [1] to derive a modified Cramér-Rao bound (MCRB) for dual-polarization M-QAM systems, under the assumption of low to moderate baud rates. Since the derivation of an exact closed-form MCRB is hard, we provide a lower bound on the MCRB. Furthermore, the bound is compared to the estimation error variance of a conventional maximum likelihood-based estimator. The bound indicates that SPM and NLPN in themselves do not degrade the performance of clock recovery, but when the signal is filtered with a too narrow filter, reliable clock recovery becomes impossible.

Notation: We will denote column-vectors in bold (e.g., \mathbf{x}), matrices in bold capitals (e.g., \mathbf{X}), the conjugate transpose of \mathbf{x} by \mathbf{x}^H , the Euclidean norm of \mathbf{x} by $\|\mathbf{x}\|$, the real part of \mathbf{x} by $\Re\{\mathbf{x}\}$, the expectation operator by $\mathbb{E}\{\cdot\}$, the phase of a complex number x by $\angle x$.

II. COHERENT OPTICAL COMMUNICATION MODEL

A. Channel Model

We consider a dual-polarization M-QAM coherent optical communication system as shown in Fig. 1. The transmission system consists of multiple consecutive optical spans. In a

practical setting, each span consists of a single-mode fiber (SMF), a dispersion-compensating fiber (DCF), and an amplifier. The DCF is designed to compensate perfectly for the chromatic dispersion in the SMF. By assuming the use of DCF and restricting the symbol rate to moderate values, we can neglect chromatic dispersion. Each optical span ends with an amplifier that compensates the propagation power loss perfectly. At the receiver side, the optical signal is converted to an electrical signal, filtered, and sampled. For easy reference, the system parameters we use in this paper are given in Table I. The notations will be clarified in later sections.

B. Validity of the Model

Our model includes SPM but neglects chromatic dispersion and interchannel effects, resulting in the same model as used previously, e.g., for studying nonlinear phase noise [7], [11]–[13]. Here, we discuss this model and quantify its validity.

1) *Dispersion*: In absence of dispersion, the nonlinear phase shift in one amplifier span is $\phi_{\text{SPM}} \approx \gamma P/\alpha$, where P is the initial power. This is induced over the effective length $L_{\text{eff}} = 1/\alpha \approx 17.4$ km in SMF, since the nonlinear length $L_{\text{NL}} = 1/(\gamma P_0) \gg L_{\text{eff}}$. In order to neglect dispersion, periodical (span-wise) dispersion compensation must be used, which is still the case in most commercial links. Furthermore, the dispersion-induced waveform change should be small over L_{eff} . This implies that the dispersion length $L_{\text{D}} = 1/(B^2|\beta_2|)$, where β_2 is the group-velocity dispersion coefficient, should be much larger than L_{eff} . Equality occurs when $B = \sqrt{\alpha/|\beta_2|} \approx 51$ GHz (assuming $\beta_2 = 22$ ps²/km). For the discussed channel bandwidth (14 GHz), dispersion is a weak effect.

2) *Interchannel effects*: The model is valid for systems using wavelength-division multiplexing (WDM) when the interchannel nonlinearities, cross-phase modulation (XPM) and four-wave mixing (FWM), are small compared to SPM. The XPM-induced phase shift is, with sufficiently large channel spacing, proportional to the walk-off length $L_{\text{wo}} = 1/(B|\beta_2|\Delta\omega)$, where $\Delta\omega = 2\pi\Delta\nu$ is the channel spacing [14]. The induced phase shift is approximately $\phi_{\text{XPM}} = 2\gamma\bar{P}L_{\text{wo}}$, where \bar{P} is the power of the neighboring channel. Assuming identical channel powers, we have $\phi_{\text{XPM}}/\phi_{\text{SPM}} = 2L_{\text{wo}}/L_{\text{eff}} = 2\alpha/(B|\beta_2|\Delta\omega)$. This ratio is 15% even with a channel spacing of 400 GHz, which shows that neglecting XPM is not compatible with high spectral efficiency. This shows that the XPM influence on clock recovery deserves further study.

The impact of FWM is estimated using the parametric gain $G = 1 + (\gamma P/g)^2 \sin^2(gL)$, where $g^2 = \Delta\beta(\Delta\beta/4 + \gamma P)$ and the fiber is assumed to be lossless [15]. The linear phase mismatch $\Delta\beta = \beta_2\Delta\omega^2$ and at the power levels relevant for communication systems we have $g \approx |\beta_2|\Delta\omega^2/2$. The converted power is proportional to $1/g^2$ and $g_{\text{SPM}}^2/g_{\text{FWM}}^2$ must therefore be small in order to neglect FWM. Since SPM can be viewed as FWM within the channel spectrum itself, we have $g_{\text{SPM}}^2/g_{\text{FWM}}^2 \approx (B/\Delta\nu)^2$. Compared to XPM, the channel spacing demand is significantly less restrictive in this case.

Table I
SYSTEM AND CHANNEL PARAMETER VALUES

name	symbol	value, unit
nonlinearity parameter	γ	$1.2 \text{ W}^{-1}\text{km}^{-1}$
attenuation	α	0.25 dB/km
length per span	L	80 km
ASE constant	n_{sp}	1.5
bandwidth	$B, 1/T$	14 GHz
number of spans	N_{a}	22
number of symbols	N_{d}	100
wavelength	$\lambda = c/\nu$	1.55 μm

C. Received Signal

Two independent data sequences are transmitted through two orthogonal optical polarizations. We will consider the equivalent electrical complex baseband signals. The signal at time t at the output of the transmitter is given by a vector of length 2:

$$\mathbf{r}_0(t) = \sqrt{P_{\text{in}}} \sum_{n=1}^{N_{\text{d}}} \mathbf{U} \mathbf{a}_n p(t - nT - \tau) e^{j\theta}, \quad (1)$$

where P_{in} is the launch power per polarization, $\mathbf{a}_n = [a_n^{(\text{X})} \ a_n^{(\text{Y})}]^T$ is the n -th data symbol vector, N_{d} is the number of data symbols, $p(t)$ is a unit energy pulse, T is the symbol duration, τ is the unknown timing offset between transmitter and receiver, θ is an unknown carrier phase offset, and \mathbf{U} is an unknown unitary matrix that models the mixing of the two polarization-multiplexed channels that occurs during transmission. We consider conventional return-to-zero (RZ) pulses [16] with 33% and 67% duty cycles, given in [17]. We assume that $\mathbb{E}\{\mathbf{a}_n\} = \mathbf{0}$, and $\mathbb{E}\{\mathbf{a}_n^H \mathbf{a}_m\} = 0$, when $m \neq n$, where $\mathbb{E}\{\cdot\}$ denotes the expectation operator. The optical signal after the i -th amplifier is given recursively by

$$\mathbf{r}_i(t) = \mathbf{r}_{i-1}(t) \exp\left(j\gamma L_{\text{eff}} \|\mathbf{r}_{i-1}(t)\|^2\right) + \mathbf{n}_i(t), \quad (2)$$

where γ is the nonlinearity parameter of the SMF, and $L_{\text{eff}} = (1 - e^{-\alpha L})/\alpha$ is called the effective length of the SMF for attenuation α and physical length L . The noise due to amplified spontaneous emission (ASE) $\mathbf{n}_i(t)$ is modeled as a circularly symmetric complex Gaussian noise process with power spectral density $N_0 = h\nu n_{\text{sp}}(G - 1)$ in each polarization, where h is Planck's constant, ν is the optical frequency, n_{sp} is the spontaneous emission factor and G is the power gain of the amplifier. Substituting $\mathbf{r}_{i-1}(t)$ recursively, we find that the received signal $\mathbf{r}(t) \triangleq \mathbf{r}_{N_{\text{a}}}(t)$ can be expressed as

$$\mathbf{r}(t) = \mathbf{r}_0(t) \exp\left(j\gamma L_{\text{eff}} \sum_{i=0}^{N_{\text{a}}-1} \|\mathbf{r}_i(t)\|^2\right) + \mathbf{w}(t), \quad (3)$$

where $\mathbf{w}(t)$ is a complex white Gaussian noise process with power spectral density $N_{\text{a}}N_0$ in each polarization. The goal of clock recovery is to determine τ , without knowledge of \mathbf{a} , θ , or \mathbf{U} . In practice, clock recovery generally takes place after chromatic dispersion compensation [18].

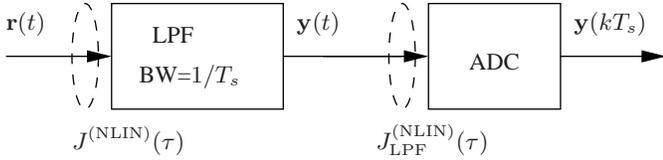


Figure 2. The final received signal $\mathbf{r}(t)$ is filtered by an ideally flat low-pass filter (LPF) bandwidth equal to $1/T_s$ and sampled, using an analog-to-digital converter (ADC), at rate $1/T_s$. We compute the Fisher information $J^{(\text{NLIN})}(\tau)$ before the LPF and the Fisher information $J_{\text{LPF}}^{(\text{NLIN})}(\tau)$ after the LPF.

III. MCRB FOR CLOCK RECOVERY

A. Background

The MCRB is a tool from estimation theory [10] that establishes a (loose) lower bound on the variance of any unbiased¹ estimator, and is thus useful to understand ultimate performance limits and to assess the performance of practical estimators in an absolute frame of reference. Given a vector representation \mathbf{r}_{vec} (obtained, e.g., by sampling at a sufficiently high rate) corresponding to an observed signal and an unknown deterministic parameter τ , the MCRB is defined as the inverse of the so-called *Fisher information*, given by

$$J(\tau) = -\mathbb{E} \left\{ \frac{\partial^2 \ln p(\mathbf{r}_{\text{vec}}|\tau, \mathbf{a}, \theta, \mathbf{U})}{\partial \tau^2} \right\}, \quad (4)$$

where $p(\mathbf{r}_{\text{vec}}|\tau, \mathbf{a}, \theta, \mathbf{U})$ is the likelihood function, and \mathbf{a} denotes the totality of transmitted symbols. To lighten the notation, we do not write $J(\tau)$ as a function of θ or \mathbf{U} , for reasons that will become clear later. The expectation in (4) is over the noise and the unknown data symbols. The MCRB is $\text{MCRB}(\tau) = J^{-1}(\tau)$, and has the following property: for any unbiased estimator $\hat{\tau}$, the variance of the estimation error obeys the following inequality:

$$\text{var}(\tau - \hat{\tau}) \geq \text{MCRB}(\tau). \quad (5)$$

We will derive an expression for the likelihood function $p(\mathbf{r}_{\text{vec}}|\tau, \mathbf{a}, \theta, \mathbf{U})$, and then determine the MCRB for a simplified case, with and without low-pass filter (LPF), shown in Fig. 2.

B. The Likelihood Function of the Received Signal

We will consider two models, the first one corresponding to Fig. 1, and a simplified model that allows closed-form computation of the MCRB.

1) *Model I*: To obtain a vector representation of the received signal, we will assume that the signal has a bandwidth W (i.e., the received signal outside of $[-W/2, W/2]$ contains only additive noise). Then, \mathbf{r}_{vec} is obtained by filtering $\mathbf{r}(t)$ with a filter that has a frequency response equal to 1 for $f \in [-W/2, +W/2]$ and is zero for all other frequencies, followed by sampling at a rate W . Note that W depends on the input power, as the signal will experience spectral broadening.

¹An estimator is said to be unbiased when the expected value of the estimate is equal to the true value of the unknown parameter.

We can substitute (2) into (3) and obtain the sufficient statistics [19]

$$\mathbf{r}_k = \mathbf{s}_k \exp \left(j\gamma L_{\text{eff}} N_a \|\mathbf{s}_k\|^2 + j\delta_k + j\phi_k \right) + \mathbf{w}_k, \quad (6)$$

where $\mathbf{r}_k = \mathbf{r}(k/W)$, $\mathbf{s}_k = \mathbf{s}(k/W) = \mathbf{r}_0(k/W)$, $\mathbf{w}_k \sim \mathcal{CN}(\mathbf{0}, N_a \sigma_{\text{ase}}^2 \mathbf{I})$, $\sigma_{\text{ase}}^2 = N_0 W$. Here, $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a multi-variate complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The angles δ_k and ϕ_k are distributed as follows: conditioned on \mathbf{s}_k , δ_k is a zero-mean Gaussian random variable with variance

$$\sigma_{\delta_k}^2 = \gamma^2 L_{\text{eff}}^2 \|\mathbf{s}_k\|^2 (N_a - 1) N_a (2N_a - 1) \sigma_{\text{ase}}^2 / 6, \quad (7)$$

and ϕ_k is a χ^2 random variable, independent of \mathbf{s}_k . We can neglect ϕ_k when $P_{\text{in}} \gg \sigma_{\text{ase}}^2$.

2) *Model II*: As Model I is hard to deal with analytically, due to the dependence of δ_k on the data and the noise, we will consider a simplified model, where we neglect δ_k and focus on the effect of nonlinear phase shift caused by the original signal. We will revisit Model I when we consider the performance of practical estimators in Section IV. Letting \mathbf{r}_{vec} denote the sequence of the samples of the received signal, the likelihood function for this simplified model is given by

$$p(\mathbf{r}_{\text{vec}}|\tau, \mathbf{a}, \theta, \mathbf{U}) = \prod_{k=-\infty}^{+\infty} p(\mathbf{r}_k|\tau, \mathbf{a}, \theta, \mathbf{U}) \quad (8)$$

$$\propto \prod_{k=-\infty}^{+\infty} \exp \left(-\frac{\|\mathbf{r}_k - \tilde{\mathbf{s}}_k\|^2}{N_a \sigma_{\text{ase}}^2} \right) \quad (9)$$

$$\propto \prod_{k=-\infty}^{+\infty} \exp \left(-\frac{\|\tilde{\mathbf{s}}_k\|^2}{N_a \sigma_{\text{ase}}^2} \right) \exp \left(\frac{2\Re \{ \mathbf{r}_k^H \tilde{\mathbf{s}}_k \}}{N_a \sigma_{\text{ase}}^2} \right), \quad (10)$$

where we have introduced $\tilde{\mathbf{s}}_k \triangleq \mathbf{s}_k \exp \left(j\gamma L_{\text{eff}} N_a \|\mathbf{s}_k\|^2 \right)$. Note that \mathbf{s}_k depends on τ , θ , and \mathbf{U} . We note that while Model II is not interesting from a practical point of view,² the MCRB corresponding to Model II will be lower than the MCRB corresponding to Model I, as Model II has fewer noise sources. Hence, the MCRB for Model II is a valid lower bound on the variance of an estimator for Model I.

C. MCRB Derivation for Model II

1) *Case 1: without low-pass filter*: The log-likelihood function can be expressed as

$$\ln p(\mathbf{r}_{\text{vec}}|\tau, \mathbf{a}, \theta, \mathbf{U}) \propto \frac{1}{N_a \sigma_{\text{ase}}^2} \sum_{k=-\infty}^{+\infty} (2\Re \{ \mathbf{r}_k^H \tilde{\mathbf{s}}_k \} - \tilde{\mathbf{s}}_k^H \tilde{\mathbf{s}}_k) \quad (11)$$

where \propto denotes proportionality up to additive constants, independent of τ . After some straightforward manipulations,

²Indeed, the nonlinearity could be precompensated at the transmitter.

we obtain

$$\begin{aligned} & \frac{\partial^2 \ln p(\mathbf{r}_{\text{vec}}|\tau, \mathbf{a}, \theta, \mathbf{U})}{\partial \tau^2} \\ &= \frac{1}{N_a \sigma_{\text{ase}}^2} \sum_{k=-\infty}^{+\infty} \left(2\Re \left\{ [\mathbf{r}_k^H - \tilde{\mathbf{s}}_k^H] \frac{\partial^2 \tilde{\mathbf{s}}_k}{\partial \tau^2} \right\} - 2 \frac{\partial \tilde{\mathbf{s}}_k^H}{\partial \tau} \frac{\partial \tilde{\mathbf{s}}_k}{\partial \tau} \right). \end{aligned} \quad (12)$$

Next, taking the expectation of (12) over the noise we get

$$J^{(\text{NLIN})}(\tau) = \frac{2}{N_a \sigma_{\text{ase}}^2} \sum_{k=-\infty}^{+\infty} \mathbb{E} \left\{ \frac{\partial \tilde{\mathbf{s}}_k^H}{\partial \tau} \frac{\partial \tilde{\mathbf{s}}_k}{\partial \tau} \right\}, \quad (13)$$

where the remaining expectation is with respect to the data symbols. The complete derivation of $J^{(\text{NLIN})}(\tau)$ is given in the Appendix, and yields

$$\begin{aligned} J^{(\text{NLIN})}(\tau) &= \underbrace{\frac{2P_{\text{in}}N_dE_2}{N_aN_0} \int_{-\infty}^{+\infty} |\dot{p}(t)|^2 dt}_{\doteq J^{(\text{LIN})}(\tau)} \\ &+ \frac{8\gamma^2L_{\text{eff}}^2P_{\text{in}}^3N_dN_aE_6}{N_0} \int_{-\infty}^{+\infty} |\dot{p}(t)|^2 |p(t)|^4 dt, \end{aligned} \quad (14)$$

where $E_l = \mathbb{E}\{\|\mathbf{a}_n\|^l\}$ and $\dot{p}(t)$ denotes differentiation of $p(t)$ with respect to t . We observe that the Fisher information comprises two terms: the first term $J^{(\text{LIN})}(\tau)$ is the well-known Fisher information for linearly modulated signals without SPM [20]. The second term is non-negative, and can be interpreted as the *additional* information due to the nonlinearity. This can be explained by the widened spectrum of the signal due to SPM. The high-frequency components provide more information regarding the parameter τ , thus increasing the Fisher information. The additional information is cubic in the input power, and will thus be more prominent at higher input powers. Note also that the Fisher information does not depend on θ , \mathbf{U} , or τ .

2) *Case 2: with low-pass filter:* In this section we assume that the received signal is further filtered with a low-pass filter $h(t)$ that has a frequency response equal to 1 for $f \in [-1/(2T_s), +1/(2T_s)]$ and is zero for all other frequencies. In contrast to the previous section, we do not assume that the filter bandwidth exceeds the signal bandwidth W . The filtered signal $\mathbf{y}(t)$ is then sampled at a rate $1/T_s$ to obtain sufficient statistics. Hence, filtering will cause a loss of information and the Fisher information $J_{\text{LPF}}^{(\text{NLIN})}(\tau)$ will depend on T_s . As $1/T_s \rightarrow W$, we expect $J_{\text{LPF}}^{(\text{NLIN})}(\tau) \rightarrow J^{(\text{NLIN})}(\tau)$. Denoting convolution by \otimes , and introducing

$$\begin{aligned} q_n(t - \tau; \mathbf{a}_n) &= \\ p(t - nT - \tau) e^{j\gamma L_{\text{eff}} P_{\text{in}} N_a \|\mathbf{a}_n\|^2 |p(t - nT - \tau)|^2} \otimes h(t), \end{aligned} \quad (15)$$

the received signal after filtering can now be written as

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{r}(t) \otimes h(t) \\ &= \sqrt{P_{\text{in}}} \sum_{n=1}^{N_d} \mathbf{U} \mathbf{a}_n q_n(t - \tau; \mathbf{a}_n) e^{j\theta} + \mathbf{w}(t) \\ &\doteq \mathbf{x}(t) + \mathbf{w}(t). \end{aligned} \quad (16)$$

(17)

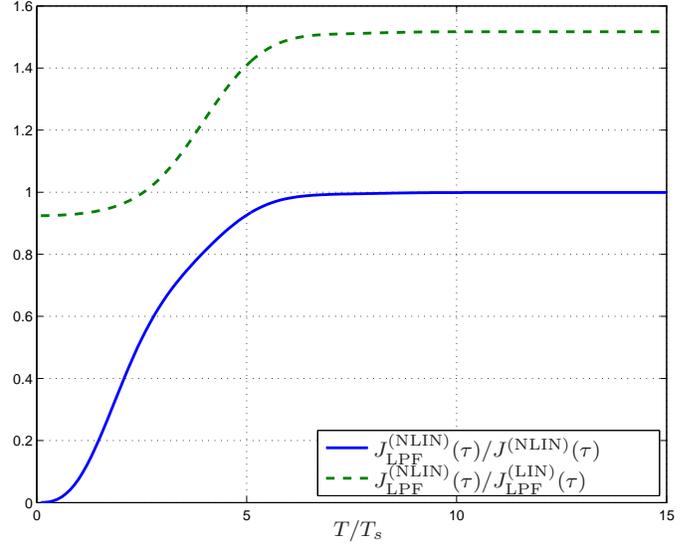


Figure 3. Impact of low-pass filter on the Fisher information as a function of oversampling factor T/T_s for $\gamma L_{\text{eff}} P_{\text{in}} N_a = 0.3$.

From $\mathbf{y}(t)$, we can compute the Fisher information $J_{\text{LPF}}^{(\text{NLIN})}(\tau)$ as

$$J_{\text{LPF}}^{(\text{NLIN})}(\tau) = \frac{2}{N_a \sigma_{\text{ase}}^2} \sum_{k=-\infty}^{+\infty} \mathbb{E} \left\{ \frac{\partial \mathbf{x}_k^H}{\partial \tau} \frac{\partial \mathbf{x}_k}{\partial \tau} \right\}, \quad (18)$$

where $\mathbf{x}_k = \mathbf{x}(kT_s)$. We easily find that

$$\begin{aligned} & J_{\text{LPF}}^{(\text{NLIN})}(\tau) \\ &= \frac{2P_{\text{in}}N_d}{N_a\sigma_{\text{ase}}^2} \mathbb{E} \left\{ \|\mathbf{a}_n\|^2 \sum_{k=-\infty}^{+\infty} |\dot{q}_n(kT_s - \tau; \mathbf{a}_n)|^2 \right\} \\ &= \frac{2P_{\text{in}}N_d}{N_aN_0} \mathbb{E} \left\{ \|\mathbf{a}_n\|^2 \int_{-\infty}^{+\infty} |\dot{q}_n(t; \mathbf{a}_n)|^2 dt \right\} \\ &= \frac{8\pi^2P_{\text{in}}N_d}{N_aN_0} \mathbb{E} \left\{ \|\mathbf{a}_n\|^2 \int_{-1/T_s}^{+1/T_s} f^2 |Q_n(f; \mathbf{a}_n)|^2 df \right\}, \end{aligned} \quad (19)$$

where $Q_n(f; \mathbf{a}_n)$ is the Fourier transform of $q_n(t; \mathbf{a}_n)$. The computation of $Q_n(f; \mathbf{a}_n)$ as well as the expectation are easily performed numerically. Again, the Fisher information does not depend on θ , \mathbf{U} , or τ . When $\gamma = 0$, we can evaluate (19) to recover the well-known Fisher information $J_{\text{LPF}}^{(\text{LIN})}(\tau)$ for linear channels. Some straightforward manipulations yield

$$\begin{aligned} & J_{\text{LPF}}^{(\text{LIN})}(\tau) = \\ & \frac{2P_{\text{in}}N_dE_2}{N_aN_0} \int_{-\infty}^{+\infty} \left| \int_{-\infty}^{+\infty} h(u) \dot{p}(t-u) du \right|^2 dt, \end{aligned} \quad (20)$$

which reverts to $J^{(\text{LIN})}(\tau)$ in (14) when $h(t) = \delta(t)$.

IV. NUMERICAL RESULTS

A. Impact of Low-Pass Filter

To gain insight into the effect of the LPF, we plot the ratios $J_{\text{LPF}}^{(\text{NLIN})}(\tau)/J^{(\text{NLIN})}(\tau)$ and $J_{\text{LPF}}^{(\text{NLIN})}(\tau)/J_{\text{LPF}}^{(\text{LIN})}(\tau)$ as a function of T/T_s for $\gamma L_{\text{eff}} P_{\text{in}} N_a = 0.3$, the pulse $p_{33}(t)$, and 4-QAM, in Fig. 3. We observe that $J_{\text{LPF}}^{(\text{NLIN})}(\tau)/J^{(\text{NLIN})}(\tau) \rightarrow 1$

as $T_s \rightarrow 0$, as we would expect. To have small information loss, we require $T/T_s > 7$, corresponding to a wideband LPF. When $T/T_s \approx 1$, which is a more practical scenario, the Fisher information is reduced by a factor 13 compared to $J_{\text{LPF}}^{(\text{NLIN})}(\tau)$. Hence, we expect significant performance degradations with a narrowband LPF. Secondly, considering $J_{\text{LPF}}^{(\text{NLIN})}(\tau)/J_{\text{LPF}}^{(\text{LIN})}(\tau)$, we observe that for low oversampling factors ($T/T_s < 2.5$) $J_{\text{LPF}}^{(\text{NLIN})}(\tau) < J_{\text{LPF}}^{(\text{LIN})}(\tau)$, meaning the nonlinearity is harmful. When the oversampling is increased, nonlinearity adds Fisher information. For $T/T_s > 7$, $J_{\text{LPF}}^{(\text{NLIN})}(\tau)$ is about 50% higher than $J_{\text{LPF}}^{(\text{LIN})}(\tau)$.

B. Impact of Input Power

We consider 4-QAM and 16-QAM signals and compare the bounds with the performance of a feed-forward (FF) timing estimator using the system parameters given in Table I. We use a conventional³ ML-based FF timing offset estimator derived for AWGN channels from [20, pp. 433-437] and extend it to dual-polarization transmission. The estimator is given as

$$\hat{\tau} = -\frac{T}{2\pi} \angle \left\{ \sum_{k=0}^{MN_d-1} [\mathbf{y}(kT_s)]^H \mathbf{z}(kT_s) e^{-j\pi k/M} \right\}, \quad (21)$$

where $M = T/T_s$, N_d is the observation length in symbols, $\mathbf{z}(kT_s) = [e^{-j\pi k/M} \mathbf{y}(kT_s)] \otimes c(kT_s)$, and $C(f) = P(f - 1/2T)P^*(f + 1/2T)$, in which $P(f)$ is the Fourier transform of $p(t)$ and $c(t)$ is the inverse Fourier transform of $C(f)$. The FF estimator has been applied to both Model I, given by (6), and Model II, given by (6), without δ_k and ϕ_k , both from Section III-B. Note that the estimator does not rely on knowledge of θ or \mathbf{U} .

In Figs. 4 and 5, $\text{MCRB}_{\text{LPF}}(\tau) = 1/J_{\text{LPF}}^{(\text{NLIN})}(\tau)$ for $T_s = T$, $\text{MCRB}(\tau) = 1/J^{(\text{NLIN})}(\tau)$, and the normalized estimation error variance with respect to input power is given for 4-QAM and 16-QAM, respectively. We observe the following:

- For $\text{MCRB}(\tau)$, we see a change in slope for higher input powers, as we expected from our discussion in Section III-C1. The impact of the constellation on $\text{MCRB}(\tau)$ is limited, as we would expect from the expression (14), where the only difference is in E_6 .
- For $\text{MCRB}_{\text{LPF}}(\tau)$, we observe that after an input power of -5 dBm, the bound stops decreasing monotonously and shows an underdamped oscillation. The effect is most obvious in 4-QAM, and becomes less pronounced for higher-order constellations. The explanation lies in oscillations in the spectrum after the LPF induced by SPM, as a function of the input power. The frequency domain oscillations cause the oscillations in the bound computed by (18). According to the results, $\text{MCRB}_{\text{LPF}}(\tau)$ depends on the type of the pulse, as well. However, the dependency is opposite to $\text{MCRB}(\tau)$: before the LPF, the bound is lower for a pulse with lower duty cycle, whereas after the LPF it is the opposite. Although the result may seem to be contradictory, it is in accordance with the expressions (14)

³Although derivation of new estimation algorithms in the presence of SPM is possible, it is not in the scope of this work.

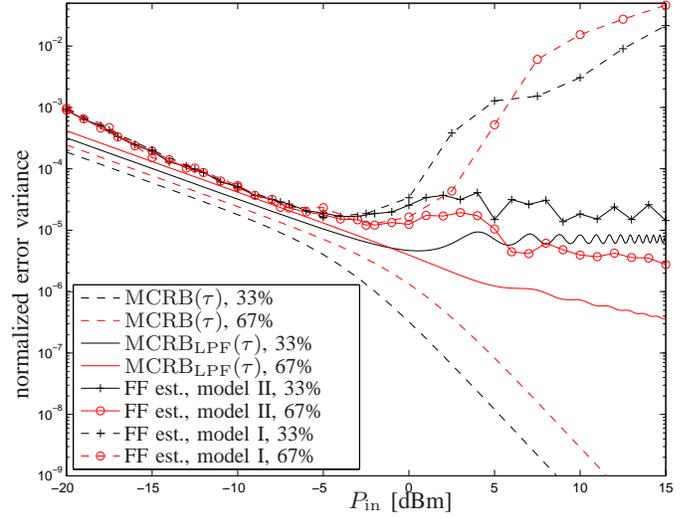


Figure 4. The MCRB with and without LPF, and normalized estimation error variance for 4-QAM for different duty cycles (33% and 66%).

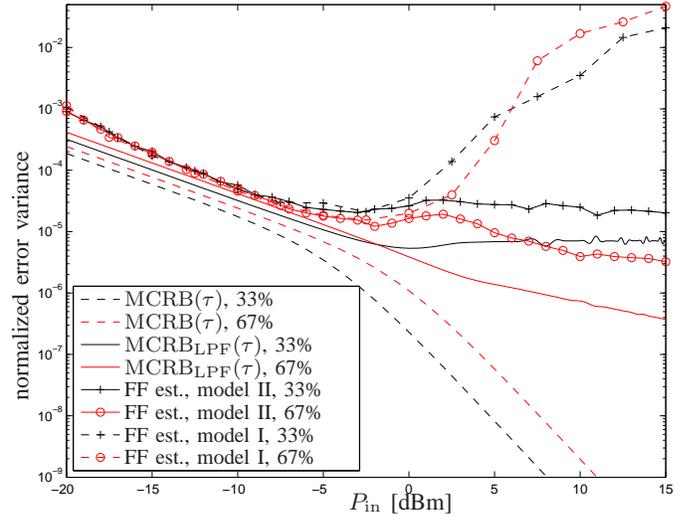


Figure 5. The MCRB with and without LPF, and normalized estimation error variance for 16-QAM for different duty cycles (33% and 66%).

and (18): the pulse with lower duty cycle has higher peak power and undergoes more spectral broadening which results in a lower bound before the LPF and a higher bound after the LPF.

- The Model II error variance follows the general behavior of $\text{MCRB}_{\text{LPF}}(\tau)$: for small-order constellations, we observe pronounced oscillations. At higher input powers, the slope of the error variance follows $\text{MCRB}_{\text{LPF}}(\tau)$.
- The Model I error variance is close to the Model II error variance for low input powers. However, beyond -5 dBm we see significant deviations, and the Model I error variance increases rapidly. This is due to the interaction between the signal and the noise, captured by the parameters δ_k in (6), which becomes more prominent at higher input powers. The behavior of the Model I error variance does not seem impacted by the constellation.

We have also observed (results not shown) that when T_s is

reduced, the bound $\text{MCRB}_{\text{LPP}}(\tau)$ is reduced, and the error floor disappears. This is congruent with our findings related to Fig. 3. Finally, we mention that the input power to reach a pre-decoding bit error rate of around 10^{-3} , which is a typical operating point, is approximately -9 dBm and -2 dBm, for 4-QAM and 16-QAM, respectively. Hence, we see that for 4-QAM, the operating point for this system corresponds to the linear regime, and traditional estimators can be used. However, for 16-QAM, the operating point corresponds to the nonlinear regime, in which case traditional estimators fail, and so the assumption of reliable clock recovery often assumed in literature may not hold when the received signal is filtered with a too narrow filter. The same holds true for higher-order modulation formats.

V. CONCLUSIONS

We have determined the MCRB for clock recovery for a coherent optical communication system with pronounced nonlinear effects (SPM and NLPN), both with and without prefilter, and compared the MCRB with the performance of practical estimators. Our analysis indicates that, for low to moderate baud rates, whereas nonlinearities are not inherently detrimental in terms of the MCRB, the use of too aggressive bandlimiting prefilters can result in significant increases in the MCRB. This implies that reliable clock recovery is impossible for high input powers when the filter at the receiver front-end is too narrow. In turn, this means that making the assumption of perfect clock recovery in this regime for the design of subsequent receiver block is meaningless, and may lead to erroneous insights and designs. Moreover, existing practical estimators, which were designed for linear transmission, exhibit poor performance in the highly nonlinear regime. Hence, (i) care should be taken when designing prefilters, and (ii) synchronization algorithms designed specifically for nonlinear transmission may result in many orders of magnitude estimation performance improvements. Future work includes the extension of the derivation to a wavelength division multiplexing setting, and the derivation of new clock recovery algorithms for the highly nonlinear regime.

ACKNOWLEDGMENT

The authors would like to gratefully acknowledge Darko Zibar for his careful reading of the manuscript and his valuable feedback.

APPENDIX

COMPUTATION OF THE FISHER INFORMATION WITHOUT LPP

Here, we will derive (14), based on the following assumptions that $p(t)$ is real, time-limited, and has support within $t \in [-T/2, +T/2]$.

Introducing $\xi_{n,k} = k/W - nT - \tau$, we can write the signal component of the received signal as

$$\begin{aligned} \tilde{\mathbf{s}}(k/W) &= \sqrt{P_{\text{in}}} \sum_{n=1}^{N_d} \mathbf{U} \mathbf{a}_n p(\xi_{n,k}) e^{j\theta} \\ &\times \exp \left(j\gamma L_{\text{eff}} N_a \left\| \sqrt{P_{\text{in}}} \sum_{n'=1}^{N_d} \mathbf{a}_{n'} p(\xi_{n',k}) \right\|^2 \right) \\ &= \sqrt{P_{\text{in}}} \sum_{n=1}^{N_d} \mathbf{U} \mathbf{a}_n p(\xi_{n,k}) e^{j\theta} \\ &\times \exp \left(j\gamma L_{\text{eff}} P_{\text{in}} N_a \|\mathbf{a}_n\|^2 |p(\xi_{n,k})|^2 \right), \end{aligned}$$

where the last equality is due to the finite duration of $p(t)$, and the fact that $\mathbf{U}^H \mathbf{U} = \mathbf{I}$. Hence

$$\begin{aligned} \frac{\partial \tilde{\mathbf{s}}_k}{\partial \tau} &= -\sqrt{P_{\text{in}}} \sum_{n=1}^{N_d} \mathbf{U} \mathbf{a}_n \dot{p}(\xi_{n,k}) e^{j\theta} \\ &\times e^{j\theta} \exp \left(j\gamma L_{\text{eff}} P_{\text{in}} N_a \|\mathbf{a}_n\|^2 |p(\xi_{n,k})|^2 \right) \\ &- 2j\gamma L_{\text{eff}} N_a P_{\text{in}}^{3/2} \sum_{n=1}^{N_d} \mathbf{U} \mathbf{a}_n \|\mathbf{a}_n\|^2 e^{j\theta} |p(\xi_{n,k})|^2 \dot{p}(\xi_{n,k}) \\ &\times \exp \left(j\gamma L_{\text{eff}} P_{\text{in}} N_a \|\mathbf{a}_n\|^2 |p(\xi_{n,k})|^2 \right). \end{aligned}$$

Then, by virtue of the assumptions on the transmitted data symbols, and the fact that $(\mathbf{U} e^{j\theta})^H (\mathbf{U} e^{j\theta}) = \mathbf{I}$,

$$\begin{aligned} \mathbb{E} \left\{ \frac{\partial \tilde{\mathbf{s}}_k^H}{\partial \tau} \frac{\partial \tilde{\mathbf{s}}_k}{\partial \tau} \right\} &= P_{\text{in}} \sum_{n=1}^{N_d} \mathbb{E} \left\{ \|\mathbf{a}_n\|^2 \right\} |\dot{p}(\xi_{n,k})|^2 \\ &+ 4\gamma^2 L_{\text{eff}}^2 N_a^2 P_{\text{in}}^3 \sum_{n=1}^{N_d} \mathbb{E} \left\{ \|\mathbf{a}_n\|^6 \right\} |p(\xi_{n,k})|^4 |\dot{p}(\xi_{n,k})|^2 \\ &+ 4\gamma L_{\text{eff}} N_a P_{\text{in}}^2 \sum_{n=1}^{N_d} \mathbb{E} \left\{ \|\mathbf{a}_n\|^4 \right\} \underbrace{\Re \left\{ j |p(\xi_{n,k})|^2 |\dot{p}(\xi_{n,k})|^2 \right\}}_{=0} \end{aligned}$$

due to the pulse being real-valued. Hence, with $E_l \doteq \mathbb{E} \left\{ \|\mathbf{a}_n\|^l \right\}$,

$$\begin{aligned} \mathbb{E} \left\{ \frac{\partial \tilde{\mathbf{s}}_k^H}{\partial \tau} \frac{\partial \tilde{\mathbf{s}}_k}{\partial \tau} \right\} &= \sum_{n=1}^{N_d} |\dot{p}(\xi_{n,k})|^2 P_{\text{in}} \left(E_2 + 4\gamma^2 L_{\text{eff}}^2 N_a^2 P_{\text{in}}^2 E_6 |p(\xi_{n,k})|^4 \right). \end{aligned}$$

Summation over all k leads to

$$\sum_{k=-\infty}^{+\infty} \mathbb{E} \left\{ \frac{\partial \tilde{s}_k^H}{\partial \tau} \frac{\partial \tilde{s}_k}{\partial \tau} \right\} = \quad (22)$$

$$WN_d P_{\text{in}} E_2 \int_{-\infty}^{+\infty} |\dot{p}(t)|^2 dt$$

$$4\gamma^2 L_{\text{eff}}^2 N_a^2 P_{\text{in}}^3 E_6 N_d W \int_{-\infty}^{+\infty} |\dot{p}(t)|^2 |p(t)|^4 dt.$$

Finally, substitution of (22) into (13) yields (14).

REFERENCES

- [1] A. S. Tan, H. Wymeersch, P. Johannisson, M. Sjödin, E. Agrell, P. A. Andrekson, and M. Karlsson, "The impact of self-phase modulation on digital clock recovery in coherent optical communication," *Proceedings ECOC*, 2010.
- [2] G. Charlet, N. Maaref, J. Renaudier, H. Mardoyan, P. Tran, and S. Bigo, "Transmission of 40 Gb/s QPSK with coherent detection over ultra-long distance improved by nonlinearity mitigation," in *Proceedings ECOC*, 2006.
- [3] R. Stolen and C. Lin, "Self-phase-modulation in silica optical fibers," *Physical Review A*, vol. 17, no. 4, pp. 1448–1453, 1978.
- [4] J. Gordon and L. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Optics letters*, vol. 15, no. 23, pp. 1351–1353, 1990.
- [5] A. P. Lau and J. M. Kahn, "Design of inline amplifier gains and spacings to minimize the phase noise in optical transmission systems," *Journal of Lightwave Technology*, vol. 24, no. 3, pp. 1334–1341, 2006.
- [6] K. P. Ho, "Mid-span compensation of nonlinear phase noise," *Optics Communications*, vol. 245, no. 1-6, pp. 391–398, 2005.
- [7] K. P. Ho and J. M. Kahn, "Electronic compensation technique to mitigate nonlinear phase noise," *Journal of Lightwave Technology*, vol. 22, no. 3, pp. 779–783, 2004.
- [8] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, 2008.
- [9] G. Li, "Recent advances in coherent optical communication," *Advances in optics and photonics*, vol. 1, no. 2, pp. 279–307, 2009.
- [10] H. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.
- [11] K. P. Ho, "Probability density of nonlinear phase noise," *J. Opt. Soc. Am. B*, vol. 20, no. 9, pp. 1875–1879, 2003.
- [12] A. P. Lau and J. M. Kahn, "Signal design and detection in presence of nonlinear phase noise," *Journal of Lightwave Technology*, vol. 25, no. 10, pp. 3008–3016, 2007.
- [13] L. Beygi, E. Agrell, M. Karlsson, and P. Johannisson, "Signal statistics in fiber-optical channels with polarization multiplexing and self-phase modulation," *J. Lightw. Technol.*, vol. 29, pp. 2379–2386, Aug. 2011.
- [14] M. Shtaif, "Analytical description of cross-phase modulation in dispersive optical fibers," *Opt. Lett.*, vol. 23, pp. 1191–1193, Aug 1998.
- [15] M. E. Marhic, N. Kagi, T.-K. Chiang, and L. G. Kazovsky, "Broadband fiber optical parametric amplifiers," *Opt. Lett.*, vol. 21, pp. 573–575, Apr 1996.
- [16] H. Zhao, E. Agrell, and M. Karlsson, "Unequal bit error probability in coherent QPSK fiber-optic systems using phase modulator based transmitters," *European Transactions on Telecommunications*, vol. 19, no. 8, pp. 895–906, 2008.
- [17] E. Ip and J. M. Kahn, "Power spectra of return-to-zero optical signals," *Journal of Lightwave Technology*, vol. 24, no. 3, pp. 1610–1618, 2006.
- [18] T. Tanimura, T. Hoshida, S. Oda, H. Nakashima, M. Yuki, Z. Tao, L. Liu, and J. Rasmussen, "Digital clock recovery algorithm for optical coherent receivers operating independent of laser frequency offset," in *34th European Conference on Optical Communication (ECOC)*, pp. 1–2, sept 2008.
- [19] A. S. Tan, H. Wymeersch, P. Johannisson, E. Agrell, P. A. Andrekson, and M. Karlsson, "An ML-based detector for optical communication in the presence of nonlinear phase noise," in *Proc. IEEE Int. Conf. on Commun.*, June 2011.
- [20] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. Plenum Press, NY, 1997.